Transient Analysis of Complex Dynamical Systems in the Context of Sustainability

Dissertation
zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach: Physik

Spezialisierung: Theoretische Physik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

MSc. Tim Kittel

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät

Prof. Dr. Elmar Kulke

Gutachter/innen:	1. Prof. Dr. Dr. h.c. mult. Jürgen Kurths
	2. Prof. Francisco Aparecido Rodrigues, MSc., PhD
	3. Dr. Sarah E. Cornell
Tag der mündlichen Prüfung:	February 7, 2018

PhD Thesis

doctor rerum naturalium (Dr. rer. nat.)

Transient Analysis of Complex Dynamical Systems in the Context of Sustainability

PhD Thesis of:
MSc. Tim Kittel

Submitted to:
Humboldt-Universität zu Berlin
Faculty of Mathematics and Natural Sciences

In coooperation with:
Potsdam Institute for Climate Impact Research
Research Domain IV: Transdisciplinary Concepts & Methods

Supervision:
Dr. Jobst Heitzig
Prof. Dr. Dr. h.c. mult. Jürgen Kurths
Prof. Francisco Aparecido Rodrigues, MSc., PhD

Submission:
for referees: **November 14, 2017**
final version: **February 19, 2018**

Referees:
1. Prof. Dr. Dr. h.c. mult. Jürgen Kurths
2. Prof. Francisco Aparecido Rodrigues, MSc., PhD
3. Dr. Sarah E. Cornell

Date of the Oral Exam:
February 7, 2018

Abstract

An important feature of dynamical systems is the transient phase of a trajectory, i.e. the part of the trajectory distant from the attractor. In this PhD thesis, I tackle two questions concerning this transient analysis: (i) "How can we properly quantify the time to reach a system's attractor?" and (ii) "Can we avoid transgressing certain boundaries and stay safe (& just)?" In particular, I consider these questions in the context of sustainability science and transformations to a sustainable future, focusing to natural and socio-economic models.

The original version of the question (i) was: "How long does a trajectory spend in the transient phase?" Observing that for most systems the mathematically precise answer is infinity leads to the reformulation stated as (i). In this thesis, I analyze several problems that come up when quantifying such transient times and define four conditions that a metric answering question (i) should fulfill. Further, I introduce two metrics, *Area under Distance Curve* (D) and *Regularized Reaching Time* (T_{RR}), capturing two complementary aspects of the transient dynamics. D is the distance of the trajectory to the attractor integrated over time and measures how "reluctant" or "eager" it is. In contrast, T_{RR} quantifies the additional time (positive or negative) a trajectory starting at a chosen initial condition needs in order to approach the attractor with comparison to some reference trajectory. It measures how much "earlier" or "later" a trajectory approaches the attractor. Further, I extend the definition of T_{RR} to networks by creating the metric *single-node recovery time* ($T_{RR}^{\mathcal{N}}$), analyzing how quickly the system can recover after a perturbation at a node. I apply them to multiple paradigmatic example and discuss in detail how they fulfill the aforementioned conditions.

Question (ii) concerns with a different point of view on the transient dynamics. Often, there are distinctions of the state space in desirable and undesirable. In particular in the context of sustainability there are such regions, e.g. defined by "planetary boundaries", "tolerable environment and development window", "guardrails", or the "safe (and just) operating space for humanity". To understand systems with such boundaries, there is not only the need to consider the intrinsic dynamics but also the possible ways of influence, i.e. managing it. I present a variant definition of the mathematical framework *Topology of Sustainable Management* (TSM) that was developed as a tool to analyze models with respect to their intrinsic dynamics, possible management and desirable states. I illustrate the concepts with multiple two-dimensional, manually-analyzed example systems. The variant definition is built on concepts from viability theory (VT). This opens the possibility to use the tools from VT, in particular the Saint-Pierre algorithm (SPA). Hence, I go the step from two-dimensional examples to higher-dimensional models by extending the application of the SPA to TSM. Furthermore, I present an extension of SPA to compute implicitly defined capture basins, a notion from VT that is more elaborated in the thesis, as these come up in TSM. I use a three-dimensional model focusing on climate change, economic output and energy transformation to demonstrate the applicability of this approach. Two common problems of estimations in VT (using SPA) are substantial for this example: (i) an unbounded state space and (ii) highly varying time scales. I solve

both by introducing appropriate, general coordinate transformations and apply them to the three-dimensional example system.

Finally, I present work that was originally devised to apply the aforementioned ideas to the El Niño-Southern Oscillation (ENSO) using *functional climate networks*. However, during the process the direction of this work deviated. I use *functional climate networks* to analyze how major perturbations – the El Niño and La Niña phases of ENSO and three largest volcanic eruptions since the middle of the 20^{th} century, Mount Pinatubo (1991), Mount Agung (1963) and El Chichon (1982) – influence the teleconnectivity structure of the surface area temperature field (SAT). The results confirm the existence of global effects of ENSO by breaking down the modular structure of the global SAT field, signifying the emergence of strong teleconnections. I find similar effects after the aforementioned volcanic eruptions on regional scales.

German Abstract

Ein wichtiger Aspekt der Analyse von dynamischen Systemen ist die Transiente einer Trajektorie. Dies ist der Teil, welcher entfernt vom Attraktor ist. In dieser Dissertation bearbeite ich diesbezüglich zwei Fragen: (i) "'Wie kann man in geeigneter Form die Zeit zum Erreichen des Attraktors quantifizieren?"' und (ii) "'Kann man es verhindern, bestimmte Grenzen zu überschreiben und somit sicher zu bleiben?"' Insbesondere werde ich diese Fragen im Kontext von Nachhaltigkeitsforschung und Transformationen zu einer nachhaltigen Zukunft untersuchen. Als Beispiele verwende ich Modelle mit natürlichen und sozio-ökonomischen Prozessen.

Die ursprüngliche Version von Frage (i) war: "'Wieviel Zeit verbringt eine Trajektorie in der transienten Phase?"' Die Beobachtung, dass die mathematisch präzise Antwort für die meisten Systeme unendlich ist, führt zu der umformulierten Version in (i). In dieser Dissertation analysiere ich mehrere Probleme, welche bei der Quantifizierung solcher transienter Zeiten auftreten, und definiere vier Bedingungen, die eine Antwort auf Frage (i) erfüllen soll. Weiterhin führe ich zwei Metriken, *Area under Distance Curve* (D) und *Regularized Reaching Time* (T_{RR}), ein, die verschiedene Aspekte der transienten Dynamik einfangen. D ist der Abstand der Trajektorie zum Attraktor integriert über die Zeit und misst wie "'reluctant"' (zurückhaltend) oder "'eager"' (eifrig) sie ist. T_{RR} hingegen quantifiziert die zusätzliche Zeit (positiv oder negativ), die eine Trajektorie im Vergleich zu einer Referenztrajektorie braucht, um sich dem Attraktor anzunähern. Diese Metrik misst, wie viel "'earlier"' (früher) oder "'later"' (später) diese Trajektorie sich annähert. T_{RR} erweitere ich noch auf Netzwerke, indem ich die Metrik *single-node recoverz time* ($T_{RR}^{\mathcal{N}}$) definiere. Sie analysiert, wie schnell sich ein System nach einer Störung an einem Knoten wieder in den Ursprungszustand begibt. Alle Metriken wende ich auf mehrere paradigmatische Beispiele an und diskutiere detailliert wie sie die vorher erarbeiteten Bedingungen erfüllen.

Frage (ii) bezieht sich auf einen anderen Aspekt der transienten Dynamik. Oft gibt es im Zustandsraum eines Systems erwünschte und unerwünschte Zustände. Insbesondere im Kontext der Nachhaltigkeit wurden solche Regionen definiert, z.B. durch "'Planetare Grenzen"'. Um Systeme mit solchen Grenzen zu verstehen, ist es notwendig, zusätzlich zur intrinsischen Dynamik auch die Möglichkeiten zur Beeinflussung, also das Management, zu betrachten. Ich werde eine Neudefinition des mathematischen Analysekonzept "'Topology of Sustainable Management"' (Topologie des nachhaltigen Managements, TSM) vorstellen, welche entwickelt wurde um Modelle in Bezug auf ihre intrinsische Dynamik, mögliches Management und erwünschte Zustände zu analysieren. Dieses Konzept wende ich auf eine Anzahl von zwei-dimensionalen, manuell analysierten Modellen an. Die Neudefinition basiert auf den Konzepten der "'Viabilitätstheorie"' (Viability Theory, VT), um den dazugehörigen Saint-Pierre Algorithmus (SPA) zu verwenden. Durch dessen Anwendung auf TSM automatisiere ich die Analyse und kann sie auf höher-dimensionalen Modelle anwenden. Dazu stelle ich eine Erweitung des Algorithmus zur Schätzung impliziter "'Capure Basins"' vor, welche innerhalb von TSM vorkommen. Als Anwendungsbeispiel nehme ich ein koevolutionäres Modell, das auf Klimawandel,

Wirtschaftsleistung und die Transformation des Energiesystems fokussiert. Zwei häufige Probleme von SPA treten hier auch auf: ein unbeschränkter Zustandsraum und stark variierende Zeitskalen. Ich stelle jeweils eine passende, allgemeine Koordinatentransformation zur Lösung vor und wende beide in dem Beispiel an.

Danach präsentiere ich eine Arbeit, bei der die schon genannten Ideen auf die El Niño-Southern Oscillation (ENSO) anwendet werden sollten. Jedoch hat sich im Bearbeitungsprozess die Richtung geändert. Ich nutze funktionale Klimanetzwerke, um zu analysieren, wie bedeutende Störungen – die El Niño- und La Niña-Phasen von ENSO und die drei größten Vulkaneruptionen seit Mitte des 20. Jahrhunderts, Mount Pinatubo (1991), Mount Agung (1963) und El Chichon (1982) – sich auf die Telekonnektionsstruktur der globalen Oberflächentemperatur (SAT) auswirken. Die Resultate bestätigen der globalen Einfluss von ENSO durch das Zusammenbrechen der modularen Struktur des global SAT-Feldes. Dies zeigt die Emergenz starker Telekonnektionen. Ähnliche, regionalere Effekte beobachte ich bei den Vulkaneruptionen.

Acknowledgements

I am grateful to my Supervisors Jobst Heitzig, who has always been open for extensive research discussions with a cup of green tea, Jürgen Kurths, who opened so many doors for me and promoted me in all aspects of my scientific life, and Francisco Rodrigues, who guided my during my stay in Brazil and throughout my thesis.

I am deeply thankful to Catrin Ciemer and Finn Müller-Hansen. They are friends, colleagues, collaborators, discussion partners and travel companions with whom I shared so many good times.

For interesting scientific interactions and collaborations, I thank Reik Donner, Chiranjit Mitra, Frank Hellmann, Rebekka Koch, Karsten Bölts, Paula Breitbach, Kevin Webster, Lyubov Tupikina, Paul Schultz, Thomas Peron, Jaap Eldering, Sabine Auer, Jasper Franke, Tiago Pereira, Stefan Ruschel, Jonathan Donges, Jakob Kolb, Wolfram Barfuss, Till Koster, Jan Nitzbon, Ilona Otto, Kilian Zimmerer, and Nastaran Lotfi

For their support during my stay in Brazil, offering me a place to stay and supporting me with my academic and private life there, I thank Roberto Gueleri, Fabiano Berardo de Sousa, Julien Korinman, Tiago Pereira, Stefan Ruschel, Edmilson Roque dos Santos, and Camille Poignard.

I thank David Hansmann for all his work throughout the years. I thank the PIK-IT-Team, particularly Dietmar Gibietz-Rheinbay, Ciaron Linstead and Benjamin Kriemann, for their support with all technical problems that came up during my PhD. I am grateful to the IRTGT 1740 “Dynamical Phenomena in Complex Networks: Fundamentals and Applications”, funded by the DFG/FAPESP, that financially enabled me to carry out the presented research.

Finally, I thank you, C., for inspiring my life and accompanying me on the way.

Contents

List of Figures

List of Tables

List of Models

List of Abbreviations

Abbreviation	Description
PBs	Planetary Boundaries
IPBs	Interacting Planetary Boundaries
SFs	social foundations
SAJOS	safe and just operating space
TSM	Topology of Sustainable Management
VT	Viability Theory
SPA	Saint-Pierre algorithm
ENSO	El Niño-Southern Oscillation
SAT	surface air temperature
RHS	right-hand-side function of an ordinary differential equation

Table 0.1: Description of abbreviations used in the thesis.

List of Symbols

The following list describes all symbols that are of importance for the entire thesis. Symbols that are only used at specific points within the thesis, e.g. parameters of an example model, are not given here.

Symbol	Description
A	adjacency matrix of a network
$\mathscr{A}$	the contextually relevant attractor of a deterministic dynamical system
$\mathscr{B}_{\mathscr{A}}$	the corresponding Basin of Attraction
bc_i	the betweenness of a node in a network
Capt $(\cdot)$	capture basin function
D	Area under Distance Curve, a metric defined in Section 3.3.1
$\mathscr{D}$	the downstream
Symbols	

Symbol	Description
$\mathscr{D}^{(+)}$	the remaining sunny downstream
$\mathscr{D}^{-}$	the dark downstream
$d(\cdot, \cdot)$	distance function, particularly used in combination with D
$\mathscr{E}$	the eddies
$\mathscr{E}^{+}$	the sunny eddies
$\mathscr{E}^{-}$	the dark eddies
f	the right-hand-side function (RHS) of an ordinary differential equation
F	RHS for the node dynamics of a network
G	RHS for the interaction between nodes of a network
$\mathscr{G}$	the glades
h	the right-hand-side function (RHS) of a dynamical control system
$\tilde{h}$	compactified RHS, see Equation (4.34)
$\tilde{\tilde{h}}$	rescaled RHS after *time homogenization*, see Section 4.5.4
k_i	the degree of a node in a network
$\mathscr{L}$	the lakes
$\mathscr{L}_u$	the time-unlimited lakes
$\mathscr{L}_l$	the time-limited lakes
$\mathscr{M}$	the manageable region
m	number of links in a network
N	number of nodes in a network
$\mathscr{N}$	set of nodes in a network
n	dimension of a dynamical system
$\mathcal{Q}$	the modularity of a network
q	solution of a dynamical (control) system
$\mathcal{S}$	the shelters
$\mathcal{T}$	the (global) transitivity of a network

Symbols

Symbol	Description
T_F^ϵ	first entry time to ϵ-neighborhood of the relevant attractor
T_L^ϵ	last entry time to ϵ-neighborhood of the relevant attractor
T_{RR}	Regularized Reaching Time, a metric defined in Section 3.3.2
$T_{RR}^{\mathcal{N}}$	single-node recovery time, a metric defined in Section 3.5
$\mathscr{u}$	general control parameter of a control system
$\mathcal{U}$	space of control parameters of a control system
$\mathscr{U}$	the upstream
$\mathscr{U}^{(+)}$	the remaining sunny upstream
$\mathscr{U}^{-}$	the dark upstream
Viab $(\cdot)$	viability kernel function
$\mathscr{W}$	the backwaters
$\mathfrak{X}$	general state space for a dynamical system
$\mathfrak{X}_i$	state space of i-th node for a networked dynamical system
$\mathfrak{X}^+$	the desirable region
$\mathfrak{X}^-$	the undesirable region
$\mathscr{x}$, $\mathscr{y}$, $\mathscr{z}$	general states for a dynamical system
$\mathscr{x}^{\mathrm{a}}$, $\mathscr{x}^{\mathrm{b}}$, ...	States with superscripts are specific/fixed states with a (contextual) meaning.
$\mathscr{x}_0$, $\mathscr{x}_1$, ...	Subscripts of states refer to the coordinate of this state. The count always starts with 0. There is an notational exception during the extension of *Timing of Transients* to networks in Section 3.5, where the subscript i refers to the i^{th} oscillator instead.
$\Delta(\cdot)$	generalized distance function used in Section 3.3.2
Θ	the trench
μ	invariant density of the attractor of interest for a dynamical system
Symbols	

Symbol	Description
ξ	distribution of initial conditions
Π	policy of a control system
ρ	distribution of perturbations
Υ	the abysses
Υ^+	the sunny abysses
Υ^-	the dark abysses
φ	time evolution operator for a dynamical system

Table 0.2: Description of the major symbols used in the thesis.

Chapter 1

Introduction

Analysis of transient behavior in physical systems, i.e. before reaching some equilibrium state or attracting set, has been done in various subfields of physics. Different phenomena during the process of magnetization for various materials, in particular the domain growth, have been studied extensively [Bar1994; Cho1979; Gau1987]. In laser physics, it was possible to derive analytical results matching the transient phases of different lasers [Fiu1980; Tan1975]. In kinetic theory, there has been research on non-equilibrium approaches for more than a century by now [Kra2010]. Other modern areas of statistical physics have emphasized the importance of transient dynamics, too, e.g. in social systems [Cas2009] and transient phases between jam and free-flow phases in vehicular traffic [Cho2000].

Recently, outside physics but within the scope of dynamical systems, an interest in transients and their dynamics has been developed. Hastings [Has2004] made a call for more transient analysis of ecological models. An example of this was given by van Geest [Van2007], describing macrophyte-dominated states of lakes as non-equilibrium states. In medicine and biology, epilepsy is seen as a transient phenomenon [Fis2005; Sch1993]. In economics, a transient analysis complementing the asymptotic analysis proved to be fruitful, particular supporting stability analysis and the understanding of how to reach the equilibria [Fis1989]. Furthermore, climate change is often seen as a transition to a new situation, i.e. a transient change to a new attractor [And2013; Len2011].

In sustainability science, understanding *transformations to a sustainable future* is a fundamental objective of research. Important key topics are the Anthropocene [Cru2002; Ste2011; Wat2016], meaning that humans have become a major geological force now, the Great Acceleration [Ste2015a], i.e. the existence of socio-economic and Earth system growth trends at rapid trends after 1950, and Planetary Boundaries (PBs) [Roc2009; Ste2015b], a set of biophysical boundaries chosen in order to ensure a sustainable human development if they are respected. From the complex dynamical systems' point of view, such a transformation can be understood as the transient behavior of a trajectory before reaching its attractor (i.e. the sustainable future).

In this thesis, I present methods for *transient analysis of complex dynamical systems*, focusing on two main points condensed in the questions:

(i) "How can we properly quantify the time to reach a system's attractor?" and

(ii) "Can we avoid transgressing boundaries and stay safe (& just)?"

I consider these questions in the context of sustainability science and the application to natural and socio-economic models.

As the work presented in this thesis has been done in cooperation with multiple people, I will use the term "we" throughout the thesis. Exceptions are information that are specific to this thesis only. In Chapter 6 'Conclusion' on page 103, I present my own contributions in more detail, using "I".

In the following I shortly present the pertinence of these questions and how they lead to the work packages of this thesis, *Timing of Transient*, *Topology of Sustainable Management*, and *Teleconnectivity Structures of the El Niño-Southern Oscillation*.

Timing of Transients. The importance of question (i) can be seen from the strong emphasis Hastings [Has2004] placed on different time scales and how the transient dynamics of a system can be rather different and more relevant than the asymptotic dynamics. In the context of sustainability science, this is demonstrated by Anderies et al. [And2013]. In this work, they present a low-complexity global carbon cycle with a saddle as a (global) desert state. Due to the nonlinearities of the system, if a trajectory starts in a certain region of the state spaces, it remains a very long time at the desert state. Hence, humanity might die out before reaching the asymptotic state, thus this long transient phase should not be neglected.

Following that thought, it seems natural to ask: "How long does a trajectory spend in the transient?" But the observation, that for most systems the mathematically precise answer is infinity, leads to the reformulation stated as (i). In the corresponding work package *Timing of Transient*, we derive and analyze the problems usually coming up during such quantification of the transient time. Based on them we define four conditions that a metric answering (i) should fulfill. We present two metrics, *Area under Distance Curve* and *Regularized Reaching Time*. The latter is extended to dynamical systems on complex networks called *single-node recovery time*. We apply these metrics to multiple paradigmatic examples and finally discuss in detail how they fulfill the aforementioned conditions.

Topology of Sustainable Management. This work package concerns question (ii). In the field of climate and sustainability sciences, boundaries have been proposed that define what a desirable situation of the Earth system is. A major concept are the *Planetary Boundaries* (PBs) introduced by Rockström et al. [Roc2009] and extended by Steffen et al. [Ste2015b]. They define a set of biophysical thresholds that we should not transgress in order to ensure human development. While they provide safety from biophysical degradation of the Earth system by defining a *safe operating space*, the social component of humans is not explicitly included. Raworth [Raw2012] pointed out the necessity to include justness (within human society) and introduced the concept of *Social Foundations* (SFs). The collection of all situations where humanity respects both of these concepts, the PBs

and SFs, is called the *safe and just operating space*, often referred to as SAJOS. While there is much research on refining the current definitions, e.g. for freshwater [Ger2013] and phosphorus [Car2011], extending them, e.g. terrestrial net primary plant production [Run2012], and downscaling them [Häy2016], another focus has emerged: their interaction due to the system's intrinsic dynamics [And2013; Hec2016], the so-called *Interacting Planetary Boundaries* (IPBs).

In this work package we present a mathematical framework of the same name, *Topology of Sustainable Management* (TSM), for the analysis of complex dynamical on a state space with a distinction in desirable and undesirable states. Further, we assume there are ways to influence the system, hence there is a *default dynamics* and *management*. The results of such an analysis is a partition of the state space in different regions, complex structures in state space corresponding to a hierarchy of safety levels. These regions are a *qualitative* classification and differ in how secure they are and how much management they need to either stay within the desirable region or reach it. Within this work certain dilemmas, i.e. a situation where a qualitative choice is present, come up. The most notable is the *lake dilemma* with a choice between (a) uninterrupted desirability (that is, not transgressing the boundaries) or (b) transgressing the boundaries temporarily but finally reaching a region where one does not have to apply management anymore. Within this thesis I present this framework first verbally and then formally, using equations based on concepts from Viability Theory (VT), a subfield of control theory. Drawing that connection enables us to use the Saint-Pierre algorithm (SPA) to automatize the model analysis. We adopt SPA for TSM and extend it for the estimation of "implicit capture basins", as these naturally arise in the TSM definition. Furthermore, two common problems of estimations in VT (using SPA) are an unbounded state space and highly varying time scales. We solve both by introducing appropriate, general coordinate transformations and apply them to the example system. Using the framework, we analyze multiple examples and show how the lake dilemma arises naturally in them.

Teleconnectivity Structures of the El Niño-Southern Oscillation. Originally, as the next step it was planned to adapt the methods developed in the first two work packages to time series analysis for real-world systems. The basic idea, motivated by Basin Stability [Men2013], was that if a time series exhibits strong perturbations in comparison to some "normal" state, these perturbations could be interpreted as "setting the initial condition" for a trajectory. Unfortunately, this did not turn out to be fruitful.

Still, we present an analysis of strong perturbations within the climate system finding global impacts and teleconnectivity structures related to the El Niño-Southern Oscillation (ENSO) as well as volcanic eruptions using *functional climate networks*. These are time-dependent networks generated by analyzing the co-variability patterns of an observable, the global surface air temperature (SAT) in our case. Using this technique, we analyze how major disruptions of the "normal" state of the global air surface temperature (SAT) affect the corresponding global teleconnectivity structure. We present an approach to quantify teleconnectivity using properties and apply it to the El Niño and La Niña phases of ENSO and three largest volcanic eruptions since the middle of the 20^{th} century, Mount

Pinatubo (1991), Mount Agung (1963) and El Chichon (1982). The results confirm the existence of global effects of ENSO by breaking down the modular structure of the global SAT field. This signifies the emergence of strong teleconnections. Furthermore, we found similar effects after the aforementioned volcanic eruptions on regional scales.

The three aforementioned work packages are presented in this thesis as follows: after this introduction, Chapter 2 gives an overview of the theoretical background for the following work. In Chapter 3, that concerns *Timing of Transients*, I go into more detail of the timing-problems, suggest two novel metrics to solve them and apply them to several examples including an extension of the concept to networks. Chapter 4 is concerned with TSM and starts with a verbal explanation of the framework, then goes into more mathematical detail by presenting the core concepts of viability theory, finally presenting a variant definition of TSM. With this definition, several 2- and 3-dimensional examples are analyzed and a bifurcation analysis of the 3-dimensional example model is included. Chapter 5 contains the aforementioned project on ENSO and volcanic eruptions. Finally, I summarize the work done and give an outlook of natural extensions of the presented work.

Where necessary, computations were done on the high performance computer system at the Potsdam Institute for Climate Impact Research.

The work presented in this thesis is based on several publications that I lead- or co-authored. These are three peer-reviewed papers [Hei2016; Kit2017f; Mit2017b] and two papers currently submitted [Kit2017d; Kit2017e]. For that reason, there is a partial overlap of the corresponding texts of the papers and this thesis. The figures in this thesis were originally published in the aforementioned papers as well. Additionally, I have three Code publications [Kit2017a; Kit2017b; Kit2017c] in order to ensure reproducibility of my results.

Chapter 2

Theoretical Background: Deterministic Dynamical Systems

In this chapter I lay down basic assumptions, notations and definitions used throughout the thesis. For more details, I refer the reader to [Per2001] and [Kuz1998], where the basics of dynamical systems are presented with all mathematical rigor. For a bigger overview, I refer the reader to [Str1994]. For an overview on networks, I recommend [New2010].

Remark. Within the thesis, I start indices, e.g. the the components of vectors or the labels of nodes in a network, from 0. In Python and C, the programming languages I used, this is commonly done, and this way I could avoid the process of shifting the index and avoid mistakes during the implementation.

2.1 Networks

Within this thesis, networks (or graphs) are are composed of N number of nodes that are linked by m links. The set of nodes is $\mathcal{N} = \{0, 1, \ldots, N-1\}$. For convenience, we use the representation by an adjacency matrix $A \in \{0,1\}^{N\times N}$. For two nodes labeled $i, j \in \mathcal{N}$, the existence of a link between them is given by $A_{ij} = 1$ and its absence by $A_{ij} = 0$. As we consider unoriented networks only, $A = A^\top$ holds and we do not allow self-loops, i.e. $A_{ii} = 0\ \forall i \in \mathcal{N}$.

Degree. The degree k_i of a node i is defined as the number of links connected to i,

$$k_i = \sum_{j\in\mathcal{N}} A_{ij}. \tag{2.1}$$

It represents how densely a node is connected within the network.

Betweenness. The betweenness bc (or sometimes betweenness centrality) of a node i is the number of shortest paths (between all other nodes) crossing i [Ant1971; Fre1977].

Put in equation, this can be formulated as

$$bc_i = \sum_{\substack{j,k \in \mathcal{N} \\ j,k \neq i}} \mathbb{1}_{\{i, j \leftrightarrow k\}}, \tag{2.2}$$

where $\mathbb{1}_{\{i,j\leftrightarrow k\}}$ is the indicator function if i lies on the shortest path between j and k. Betweenness can be understood as a measure of how central a node is within the network and might represent how influential a node is.

Transitivity. The measure characterizes how much a graph (or its sub-graphs) are clustered [Boc2006; Hei2012; Sar2007]. This is done by describing the following probability. If one chooses randomly a node and two random links, what is the probability that the two adjacent nodes are connected, too, i.e. there is a triangle? In equation, this reads as

$$\mathcal{T} = \frac{\sum_{i,j,k \in \mathcal{N}} A_{ij} A_{jk} A_{ki}}{\sum_{i,j,k \in \mathcal{N}} A_{ij} A_{jk}}. \tag{2.3}$$

Modularity. Modularity is a measure of heterogeneity within the network structure [New2004]. It describes how well different groups of nodes can be distinguished that are densely connected within each group, but only sparsely among different groups. The definition of modularity relies on a partitioning of the network into meaningful subgraphs. Here, a partition is a splitting of the nodes network in disjoint subsets of nodes, where the union of all these sets is again the full set of nodes $\mathcal{N}$. Up to a multiplicative constant, modularity counts how many links share the nodes within each subset and compares this value with the expected number of links inside if the network were random:

$$\mathcal{Q} = \frac{1}{2m} \sum_{i,j \in \mathcal{N}} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \Delta_{ij}, \tag{2.4}$$

where Δ_{ij} is an indicator function informing whether two nodes i and j belong to the same subset in the considered partition. In case of of a partition in two parts only, this equation is often written as

$$\mathcal{Q} = \frac{1}{4m} \sum_{i,j \in \mathcal{N}} \left(A_{ij} - \frac{k_i k_j}{2m} \right) s_i s_j \tag{2.5}$$

where $s_i = 1$ if i belongs to group 1 and $s_i = -1$ if i belongs to group 2 [New2006a]. In the derivation of Equation (2.5), the identity $2m = \sum_i k_i = \sum_{ij} A_{ij}$ has been used.

The individual subsets of the partition that maximizes the modularity $\mathcal{Q}$ are called *communities*. The higher the modularity, the more split-up (or modular) the network. Accordingly, community detection by modularity maximization has become a tool for cluster analysis.

2.2 Dynamical Systems

The work in thesis is mostly concerned with *deterministic, autonomous dynamical systems* with a focus on the application on *ordinary differential equations* (ODEs) as they are a common occurrence in modeling. Generally, they are given by an equation

$$\dot{x} = f(x) \qquad \forall x \in \mathcal{X} \tag{2.6}$$

with an n-dimensional state space $\mathcal{X} = \mathbb{R}^n$ and the right-hand side (RHS) $f(x)$ that we assume to be at least twice continuous differentiable. Usually, we use x to denote an arbitrary state $x \in \mathcal{X}$, and refer to specific/fixed states with letters as superscripts, e.g. x^{a}, x^{b}, and x^{ref}. The components of a state are written with subscripts, so $x = (x_0, x_1, \ldots, x_{n-1})^\top$ and $x^{\mathrm{a}} = \left(x_0^{\mathrm{a}}, x_1^{\mathrm{a}}, \ldots, x_{n-1}^{\mathrm{a}}\right)^\top$. The words "point" and "state" are used synonymously for the elements of $\mathcal{X}$.

An notational exception is used during the extension of *Timing of Transients* to networks in Section 3.5, where the subscript i refers to the i^{th} oscillator. That means $\mathcal{X}_i$ refers to the i^{th}-subspace and x_i to an element in $\mathcal{X}_i$. The reason for this exception is that both notations are straightforward and standard, hence a deviation might confuse the reader.

Time is generally denoted by t but in specific cases t and T are used to avoid possible confusions. The the set of time points is denoted by $\mathbb{T}$. If not specified differently, we assume $\mathbb{T} = \mathbb{R}_{\geq 0}$.

For convenience, we will make use of the time-evolution operator φ where $\varphi(t, x)$ is the state after starting at some point x and letting the system evolve for some time $t \geq 0$. Hence

$$\varphi(0, x) = x \quad \text{and} \tag{2.7a}$$

$$\frac{\partial \varphi}{\partial t}(t, x) = f(\varphi(t, x)). \tag{2.7b}$$

Furthermore, we will use the notion of a solution $q\colon \mathbb{T} \to \mathcal{X}$ for some initial condition x^0. That means q fulfills

$$q(0) = x^0, \tag{2.8a}$$

$$\frac{dq}{dt}(t) = f(q(t)). \tag{2.8b}$$

2.2.1 Networked Systems

If the deterministic dynamical system is based on a *network* structure, this is reflected in the system description as follows. A *networked system* with nodes $\mathcal{N}$ and a dynamics of each node in dimension ν is then a specialization of Equation (2.6) written as

$$\dot{x}_i = F(x_i) + \sum_{\substack{j \in \mathcal{N} \\ i \neq j}} A_{ij} \cdot G(x_i, x_j) \quad \forall i \in \mathcal{N}, \tag{2.9}$$

with $\boldsymbol{x}_i \in \mathbb{R}^\nu$, $F(\boldsymbol{x}_i)$ the RHS of the dynamics on the node (the oscillator), and $G(\boldsymbol{x}_i, \boldsymbol{x}_j)$ the coupling term between the oscillators. We assume F and G to be at least twice continuously differentiable and G to be symmetric, i.e. $G(\boldsymbol{x}_i, \boldsymbol{x}_j) = G(\boldsymbol{x}_j, \boldsymbol{x}_i)$ as A is symmetric as well. The ordered collection of all $\boldsymbol{x}_i$ is summarized as $\boldsymbol{x} \in \mathbb{R}^{\nu \cdot N}$.

For a bigger overview on dynamical systems on networks, I refer the reader to [New2010].

2.2.2 Attractor

For clarity, we use the basic definition of attractor found at [Wik2017] and introduce the basic types only. For mathematical details on attractors and different types, I refer the reader to [Per2001, pp. 191 sqq.].

An *attractor* $\mathcal{A} \subseteq \mathcal{X}$ of a dynamical system Equation (2.6) has three properties:

1. $\mathcal{A}$ is forward-invariant, i.e. $\varphi(t, \mathcal{A}) \subseteq \mathcal{A}\ \forall t > 0$,
2. there exists a neighborhood $\mathcal{B}_\mathcal{A} \subseteq \mathcal{X}$, called the *basin of attraction*, consisting of all points $\boldsymbol{x}$ that asymptotically approach $\mathcal{A}$, .i.e. for any neighborhood U of $\mathcal{A}$, there exists a $T > 0$ s.t. $\varphi(t, \boldsymbol{x}) \in U$ for all $t > T$, and
3. there is no non-empty proper subset of $\mathcal{A}$ that fulfills the first two properties.

We assume the system (2.6) to have at least one attractor $\mathcal{A}$ with a corresponding basin of attraction $\mathcal{B}_\mathcal{A} \subseteq \mathcal{X}$. In case the system has more than one attractor, the analysis is applied to the specified attractor of interest.

Stable equilibria. An *equilibrium* or *fixed point* is a state $\boldsymbol{x}^\mathrm{f}$ such that the RHS vanishes, i.e. $f(\boldsymbol{x}^\mathrm{f}) = 0$. We call an equilibrium which is an attractor a *stable equilibrium* or *stable fixed point*. If the derivative $Df(\boldsymbol{x}^\mathrm{f})$ of the RHS f at the point $\boldsymbol{x}^\mathrm{f}$ has full rank, then $\boldsymbol{x}^\mathrm{f}$ is called exponentially stable and locally the system can be approximated by

$$\dot{\boldsymbol{x}} = Df(\boldsymbol{x}^\mathrm{f}) \cdot \boldsymbol{x}. \tag{2.10}$$

Stable limit cycles. A *limit cycle* is a closed curve $\mathcal{P} \subset \mathcal{X}$, where all points are revisited after a fixed time $T > 0$, the *period*, i.e.

$$\varphi(T, \boldsymbol{x}) = \boldsymbol{x} \quad \forall \boldsymbol{x} \in \mathcal{P}. \tag{2.11}$$

If a limit cycle is an attractor, we call it a *stable limit cycle*.

Invariant Density. An *invariant density* is a integrable function $\mu \colon \mathcal{X} \to \mathbb{R}_{\geq 0}$ that is invariant under time evolution. In formulas, this means that for any integrable function $f \colon \mathcal{X} \to \mathbb{R}$ and for all $t > 0$ holds

$$\int_\mathcal{X} \mathrm{d}\boldsymbol{x}\ f(\boldsymbol{x})\mu(\boldsymbol{x}) = \int_\mathcal{X} \mathrm{d}\boldsymbol{x}\ f(\varphi(t, \boldsymbol{x}))\mu(\boldsymbol{x}). \tag{2.12}$$

In [Cho2006], from where I adapted Equation (2.12), they are treated as *invariant measures* (in the sense of the mathematical measure theory), as this enables him to ensure

mathematical precision. Within this thesis, we treat them as density functions and, in common physical manner, extensions like the Dirac delta function are generally included as well. Furthermore, we usually assume μ to be normalized, i.e. $\int_{\mathcal{X}} \mathrm{d}\boldsymbol{x}\ \mu = 1$.

Lyapunov exponents. The *Lyapunov exponents* capture how strongly an ensemble of trajectories in the neighborhood around the attractor is being contracted when evolving in time. Illustrative explanations and definitions can be found in [Str1994]. Most relevant for this thesis is the case for an exponentially stable equilibrium $\boldsymbol{x}^{\mathrm{f}}$. Locally, the system can then be approximated by Equation (2.10), so the system can be solved analytically by

$$\varphi(t, \boldsymbol{x}) = \sum_i a_{i,\boldsymbol{x}} v_i e^{\lambda_i t}, \tag{2.13}$$

where λ_i are the eigenvalues* of $Df(\boldsymbol{x}^{\mathrm{f}})$ with the corresponding eigenvectors v_i and $a_{i,\boldsymbol{x}}$ are coefficients that depend on the initial condition only. Hence, locally the system contracts exponentially with the rates $-\mathrm{Re}\,(\lambda_i)$ in the direction of the corresponding eigenvectors. The largest eigenvalue, that is the one with the smallest absolute value $|\mathrm{Re}\,(\lambda_i)\,|$, is called the maximal Lyapunov exponent and gives the time scale for approaching the equilibrium point.

2.2.3 Chaos

As Strogatz states "No definition of the term chaos is universally accepted yet, but almost everyone would agree on three ingredients" [Str1994, p. 323] and this is true 23 years later. The three ingredients he refers to are

1. The trajectories exhibit *aperiodic long-term behavior*, meaning they do not settle down to fixed points, periodic orbits or quasiperiodic orbits†.
2. The system is *deterministic*, i.e. there is no random or noisy input. As the whole thesis concerns with deterministic systems only, this is a given.
3. *Sensitive dependence on initial condition* means that two trajectories starting nearby separate exponentially, implying a positive Lyapunov Exponent.

A common example is the Rössler Oscillator as described in Model Description 1. For illustration of the above concepts, an estimation of its invariant density is depicted in Figures 2.1 and 2.2.

Model Description 1 (Rössler Oscillator)**.** The Rössler system is exemplary for

*Note that for a exponentially stable equilibrium all eigenvalues are negative.

†Quasiperiodicity is basically the superposition of two (or more) periodic movements with irrational quotients of periodicity. An example is given in [Str1994, pp. 273–276].

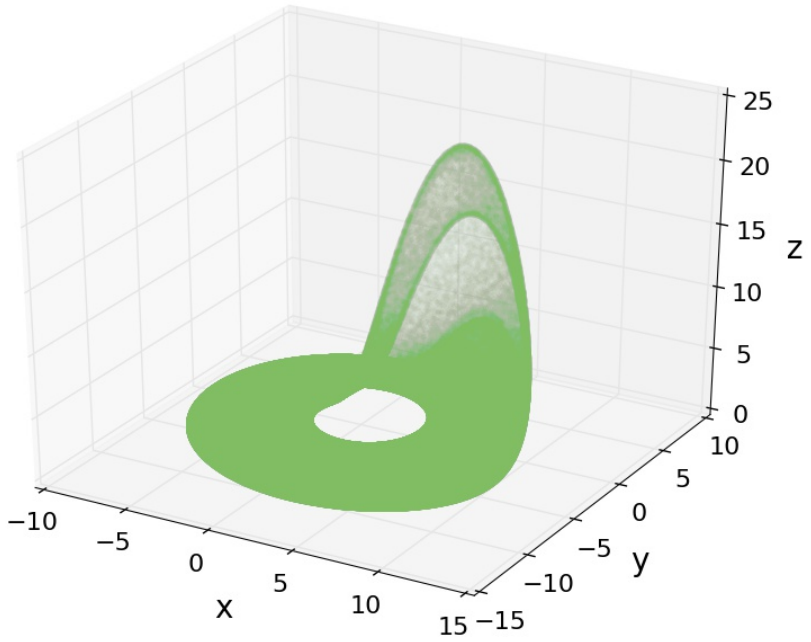


Figure 2.1: 3D-plot of the chaotic Rössler attractor's invariant density. The estimation has been done by starting at multiple initial conditions, numerically integrating the system for a long term, removing the first part of each trajectory and then plotting all trajectories with very low opacity. This graphic was published in the supplementary data of [Kit2017f].

continuous chaos in low dimensions [Rös1976; Zgl1997]. The equations are

$$\begin{aligned} \dot{x} &= -y - z, \\ \dot{y} &= x + ay, \\ \dot{z} &= b + z\,(x - c)\,, \end{aligned} \tag{2.14}$$

where x, y and z are the coordinates in state space and the state vector $\boldsymbol{x} = (x, y, z)^\top$. The standard values for the parameters are $a = 0.2$, $b = 0.2$ and $c = 5.7$ where the system exhibits a chaotic attractor. A 3-dimensional plot of the (invariant density of the) attractor is depicted in Figure 2.1 and its projections are in Figure 2.2.

2.3 Control Systems & Viability Theory

In this section, we shortly introduce the parts of *viability theory* (VT) needed in this thesis, following [Aub1991; Aub2001]. We start with a dynamical control system in the state space $\mathfrak{X} = \mathbb{R}^n$

$$\dot{\boldsymbol{x}} = h\,(\boldsymbol{x}, \boldsymbol{u}) \tag{2.15}$$

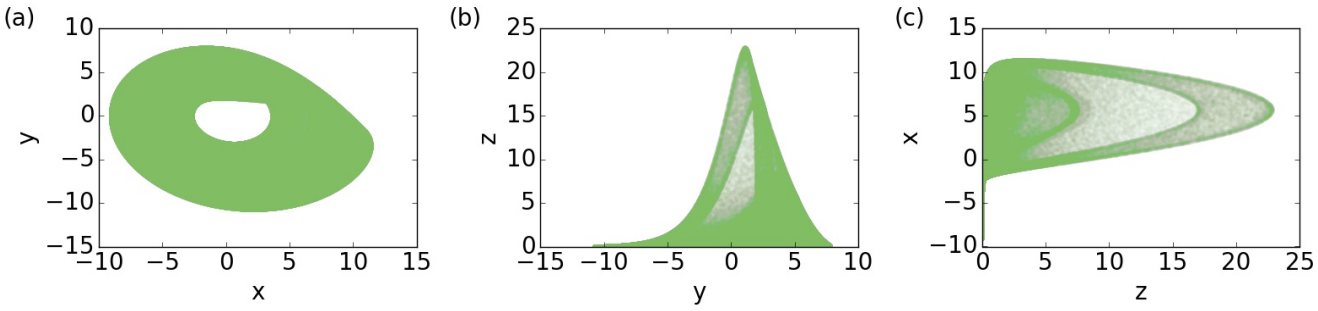


Figure 2.2: The three projections of the chaotic Rössler attractor's invariant density. The density has been obtained in the same manner as for Figure 2.1. These graphics were published in the supplementary data of [Kit2017f].

with $x \in \mathcal{X}$ and $u \in \mathcal{U}$ the set of all possible values for the control parameter u *. We call h the right-hand side function (RHS) of the control system.

A function $q\colon \mathbb{T} \to \mathcal{X}$ is called a *solution* for an (arbitrary) initial condition $x^0 \in \mathcal{X}$ if there exists a function $\Pi\colon \mathbb{T} \to \mathcal{U}$ such that for any time $t \in \mathbb{T}$ the condition $\frac{dq}{dt}(t) = h(q(t), \Pi(t))$ is fulfilled and $q(0) = x^0$. This function Π is called a *policy*.

2.3.1 Viability Kernel

The *viability kernel* of a *constraint set* $\mathcal{Y} \subseteq \mathcal{X}$ is then defined as the set of initial conditions for which there exists a *viable solution* that stays within $\mathcal{Y}$ forever

$$\begin{aligned}\mathrm{Viab}_{\mathcal{U}}(\mathcal{Y}) = \{ y^0 \in \mathcal{Y} \mid \exists\, \text{solution}\, q\colon \\ q(0) = y^0 \wedge \forall t \in \mathbb{T}\colon q(t) \in \mathcal{Y} \}\,.\end{aligned} \tag{2.16}$$

The set of possible controls $\mathcal{U}$ is given as a subscript as we will distinguish different controls. Note that the eq. (2.16) depends implicitly on the choice of $\mathcal{U}$, as the solution q is chosen with respect to some policy Π. And the latter is a function that maps from $\mathbb{T}$ to $\mathcal{U}$.

2.3.2 Capture Basin

The *capture basin* of a *target set* $\mathfrak{T} \subseteq \mathcal{X}$ is the part of the state space in $\mathcal{X}$ for which there exists a solution of Equation (2.15) that reaches $\mathfrak{T}$ in finite time while staying inside a constraint set $\mathcal{Y}$

$$\begin{aligned}\mathrm{Capt}_{\mathcal{U}}^{\mathcal{Y}}(\mathfrak{T}) = \{ x^0 \in \mathcal{X} \mid \exists \text{ solution}\, q\ \exists T \in \mathbb{T}\colon \\ q(0) = x^0 \wedge q(T) \in \mathfrak{T} \wedge \forall t < T\colon q(t) \in \mathcal{Y} \}\,.\end{aligned} \tag{2.17}$$

*Note that no dependency of $\mathcal{U}$ on the state x is assumed for notational simplicity but the whole idea could be extended in a straightforward manner.

In case no constraint set is given, the whole state space is assumed, i.e. $\mathrm{Capt}_{\mathcal{U}}(\mathfrak{X}) := \mathrm{Capt}_{\mathcal{U}}^{\mathcal{X}}(\mathfrak{X})$. On a side note, if $\mathfrak{X}$ is a *viability domain*, defined by the condition $\mathfrak{X} = \mathrm{Viab}_{\mathcal{U}}(\mathfrak{X})$, within $\mathcal{Y}$, then

$$\mathrm{Capt}_{\mathcal{U}}^{\mathcal{Y}}(\mathfrak{X}) \subseteq \mathrm{Viab}_{\mathcal{U}}(\mathcal{Y}), \tag{2.18}$$

too, implying all initial conditions within the capture basin have viable solutions.

Within the thesis, we use the term *implicit capture basins*. This refers to capture basins, that are (functions of) target sets of each other. Take $\mathcal{U}^{\mathrm{a}}$ and $\mathcal{U}^{\mathrm{b}}$ as set of control parameters and f^{a}, $f^{\mathrm{a}}\colon \mathrm{Pow}(\mathcal{X}) \to \mathrm{Pow}(\mathcal{X})$ that map between subsets of $\mathcal{X}$ *, a general form for some implicit capture basins C^{a} and C^{b} is

$$C^{\mathrm{a}} = f^{\mathrm{a}}\left(\mathrm{Capt}_{\mathcal{U}^{\mathrm{a}}}\left(C^{\mathrm{b}}\right)\right), \tag{2.19a}$$

$$C^{\mathrm{b}} = f^{\mathrm{b}}\left(\mathrm{Capt}_{\mathcal{U}^{\mathrm{b}}}\left(C^{\mathrm{a}}\right)\right). \tag{2.19b}$$

*$\mathrm{Pow}(\mathcal{X})$ denotes the power set of $\mathcal{X}$, i.e. the set of all subsets.

Chapter 3

Timing of Transients: Quantifying Reaching Times and Transient Behavior

As the main subject of the thesis is the analysis of the transient phase in deterministic dynamical systems, it is intuitive to follow up on the emphasis on *long* transients being a relevant feature for understanding dynamical, ecological systems [And2013; Has2004; Kan2016] by asking the question: "How long does a trajectory spend in the transient?" Or conversely from a more complex-system perspective: "How long does it take for a trajectory to reach the (corresponding) attractor?"

A major problem, presented in more detail in Section 3.2, is that the precise answer to these questions is infinite for a large class of attractors of physically relevant systems. So a reformulation leads to this chapter's overarching question: *"How can we properly quantify the time to reach a system's attractor?"*

Often, the division of a trajectory in transient and asymptotic part is done by arbitrarily defining when it is "close to the attractor". As this demands an analysis, we split the overarching question in two sub-questions: (a) "What are the problems of these current/intuitive methods to quantify transient time?" and (b) "How can we solve them?".

To answer the first question, we analyze the problems and consequently work out four essential conditions that we demand of a solution to the overarching question to fulfill. The attractor is not reached in finite time for a large class of physically relevant systems, so we emphasize that any such metric should be finite with *Condition I* (on page 15). As it is generally unclear how to define precisely "when the transient is over", it is ambiguous where to divide between the transient and the asymptotics. Hence, we demand a condition that relates to the physical meaning of the transient time, *Condition II*. When having parameter dependence, small changes in the parameter often induce large (non-continuous) effects on the measured quantity. Thus, in *Condition III* we demand a sensible metric to be continous. And finally, any result should not depend on the choice of coordinates, which is encoded in *Condition IV*. The latter is particularly important, as a dependence on the coordinates would imply that we measure not a systemic property but an artifact of the representation.

Then we approach question (b) by formulating two metrics, *Area under Distance*

Curve (D) and *Regularized Reaching Time* (T_{RR}). The first one is the integral over the distance to the attractor along the trajectory, and has a physical dimension of time multiplied by distance. It measures which trajectories are *reluctant*, i.e. stay distant from the attractor for long, or *eager*, i.e. approach it right away. The second one, T_{RR}, is defined by the difference between the reaching times for the trajectory of interest and a reference trajectory. Thus, it takes a different point of view, actually measuring a time. The idea is that even though the actual reaching times are infinite, their difference is typically finite. So, we can compare trajectories approaching the attractor and define the notions *earlier* and *later*.

We choose four examples to illustrate different features of these metrics. First, we use a linear system to understand how the metrics act generally and to observe the divergence of T_{RR} on the strong stable manifold particularly. Also, due to the system's simplicity, analytical solutions are possible. We then use a global carbon cycle model [And2013] and a model of a generator in the power grid [Wec2013] to apply the ideas to some first real world systems. With our final example, the chaotic Rössler oscillator, we demonstrate that we can apply these methods to more complex attractors as well. The chosen examples are rather well-understood. So they are good testing cases for the metrics, while their complexity still needs numerical approaches for a proper quantification of reaching times. This is appropriate as this work is meant as a methodological step towards applications in real-world systems. For that reason we focus on applicability with numerical estimations, and analytical results are given for a better general understanding only.

After we analyzed our two metrics for low-dimensional systems, our next step is to develop a network metric. This is done under the title *Timing on Networks* where we develop a metric that associates to each node a regularized reaching time $T_{RR}^{\mathcal{N}}$. It describes the average additional time (positive or negative) needed for a trajectory to return to the attractor after a perturbation at a particular node in comparison to a reference node. This metric is then tested on two exemplary dynamical, networked systems: coupled Rössler oscillators and the coupled Swing equation. They are applied to different topologies and analyzed: the deterministic scale-free network and the random scale-free network as well as the topology of the United Kingdom's power grid.

The rest of this chapter is structured as follows. In Section 3.2, we state the four general conditions that we expect such a metric to fulfill. They stem from problems generally encountered when dealing with the aforementioned questions and are demonstrated right after the statement of the conditions. In Section 3.3, we present the two metrics and apply them to multiple examples in Section 3.4. Afterwards, we extend the aforementioned metrics to network metrics in Section 3.5 and demonstrate them with several examples in Section 3.5.3. Finally, in Section 3.6 we discuss in detail how far the metrics fulfill the conditions developed in Section 3.2 and summarize this chapter's results in Section 3.7.

This chapter is based on the publications "Timing of transients: Quantifying reaching times and transient behavior in complex systems" [Kit2017f] and "Recovery time after localized perturbations in complex dynamical networks"[Mit2017b], where I lead-authored the first work and co-authored the second.

3.1 Wording

On a technical level, when we speak of "quantifying the transient time" we mean to find a function $\mathfrak{X} \longrightarrow \mathbb{R}$, a "metric", that gives a reasonable number for the time a trajectory spends in the transient phase for each initial condition $x \in \mathscr{B}_{\mathscr{A}}$. By taking proper statistics of this function, e.g. its average or standard deviation within $\mathscr{B}_{\mathscr{A}}$, such a metric can be used to make statements about the dynamical system. In case of networks, we refer to "network metrics", meaning a function $\mathcal{N} \longrightarrow \mathbb{R}$, that associates to each node a reasonable number. Additionally, within this chapter we assume the asymptotics of the system to be understood as we want to focus on the transient only.

3.2 Conditions on Reaching Time Definitions

In this section, we state and elaborate the four essential conditions, that we demand from metrics that quantify transient time. Afterwards, we demonstrate their necessity with an example system in order to provide the reader with a detailed understanding.

(Cond. I) *Finite Reaching Time.* The basic problem is that the time a trajectory needs to reach an attractor is typically infinity. This includes most commonly occurring steady states or limit cycles and most ODEs. As a consequence, we demand the values of a proper reaching time metric to be finite.

(Cond. II) *Physical Interpretation.* It is far from obvious what "close to the attractor" or "when the transient is over" mean precisely. Often, one addresses this by using an arbitrary threshold quantified by an ϵ to define a "small distance" to $\mathscr{A}$. But as mentioned in the description for Condition I, the trajectory might reach the attractor only in infinite time, meaning for $\epsilon \to 0$, the value diverges. So the result is strongly dependent on the parameter ϵ and this kind of dependence should be avoided. Note that the focus of this chapter is to *quantify* the transient time, i.e. the result should be a meaningful number.

(Cond. III) *Continuity.* If one is defining a metric using a parameter, there might be a discontinuous dependence on that parameter. Normally, we want these changes to be smooth and, if possible, weakly dependent on the parameter. Any discontinuous behavior should represent a corresponding property of the system that introduces this change.

(Cond. IV) *Invariance.* The metric should be a general property of the system and not depend on the chosen coordinates or variables to represent it. Changing the coordinates is simply looking at the same system from a different "point of view". For a metric to be sensible, we want it to be a genuine property of the system and independent of this "point of view", hence this condition. In mathematical terms, we demand invariance under smooth changes of coordinates (cf. [Kuz1998, p. 42]).

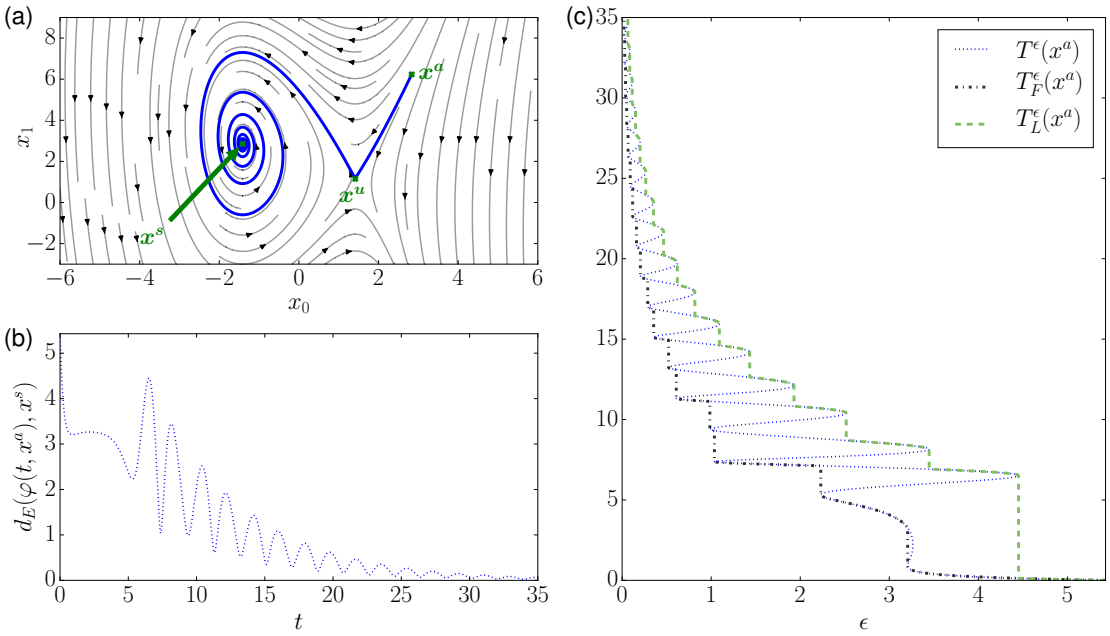


Figure 3.1: (a) shows the flow in the phase space of the example system Equation (3.1) and the stable spiraling node x^{s} and the saddle x^{u} are marked. The trajectory starting at x^{a} closely passes by x^{u} before it spirals in to x^{s}. (b) depicts the Euclidean Distance $d_E(\cdot, \cdot)$ over time for this trajectory. The first longer dip between $t = 1$ and $t = 5$ is the transient at the (unstable) saddle x^{u} while the oscillations afterwards are the spiraling around x^{s}. (c) turns (b) around in order to show the dependence of the time t on some distance $d = \epsilon$ (dotted, blue) of the trajectory to the attractor x^{s}. Secondly, there are multiple values of ϵ for each t so observables like $T^\epsilon_F(x^{\mathrm{a}})$ and $T^\epsilon_L(x^{\mathrm{a}})$ need to be introduced. $T^\epsilon_F(x^{\mathrm{a}})$ (dash-dotted, black) marks when the first time the ϵ-neighborhood around x^{s} is entered and $T^\epsilon_L(x^{\mathrm{a}})$ (dashed, green) the last time. These graphics were published in [Kit2017f].

3.2.1 Example

To demonstrate why the aforementioned Conditions I to IV are reasonable and necessary, we go through an example that has been deliberately chosen to illustrate all the problems that lead to the conditions. Still it is sufficiently simple so we do not have to cope with problems inherent to the example system.

$$\dot{x_0} = 1 - \frac{x_1}{2} - bx_0, \quad \dot{x_1} = 2\left(a - x_0^2\right), \quad \text{with} \quad a = 2, \quad b = 0.3\,. \tag{3.1}$$

The system has a stable focus at $x^{\mathrm{s}} = \left(-\sqrt{a},\ 2(1 + b\sqrt{a})\right)^\top$ as its only attractor, and an unstable saddle at $x^{\mathrm{u}} = \left(\sqrt{a},\ 2(1 - b\sqrt{a})\right)^\top$.

The flow is shown in Figure 3.1 (a). For a trajectory starting at a chosen $x^{\mathrm{a}} = (2.8, 6.2)^\top$ the Euclidean distance to the attractor over time

$$d_E\left(\varphi(t, x^{\mathrm{a}}),\, x^{\mathrm{s}}\right) = \sqrt{\varphi(t, x^{\mathrm{a}})^\top \cdot x^{\mathrm{s}}} \tag{3.2}$$

is depicted in Figure 3.1 (b). Commonly, two used metrics are the times when an ϵ-neighborhood is entered the first and the last time. Hence, we define the class of sets

$$T^{\epsilon}(\boldsymbol{x}^{\mathrm{a}}) = \{t \mid \epsilon = d_E\left(\varphi(t, \boldsymbol{x}^{\mathrm{a}}), \boldsymbol{x}^{\mathrm{s}}\right)\}, \tag{3.3}$$

that invert the axes of Figure 3.1 (b) and yield Figure 3.1 (c). Therefore, $T^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ is the set of times where an ϵ-neighborhood is entered or exited. So we find the first and last entry times to be $T_F^{\epsilon}(\boldsymbol{x}^{\mathrm{a}}) = \inf T^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ and $T_L^{\epsilon}(\boldsymbol{x}^{\mathrm{a}}) = \sup T^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ respectively, as shown in Figure 3.1 (c), too. As $T_F^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$, $T_L^{\epsilon}(\boldsymbol{x}^{\mathrm{a}}) \to \infty$ for $\epsilon \to 0$, we observe that the actual reaching time is infinite. Hence, Condition I is necessary to be mentioned for proper metric to measure the time to reach the attractor.

As both, $T_F^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ and $T_L^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$, diverge to infinity for $\epsilon \to 0$, their values depend heavily on ϵ. Therefore, a proper physical interpretation is difficult. The meaning of "close to the attractor" or "when the transient is over" depend strongly on ϵ. For that reason, we formulated Condition II.

The strong discontinuities in $T_F^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ and $T_L^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ under changes of ϵ (see Figure 3.1 (c)) increase the difficulty for a proper choice of ϵ. They arise because a trajectory enters and exits an ϵ-neighborhood (with ϵ fixed) several times, and this behavior is induced by the complex eigenvalues of the system. It could be circumvented locally by choosing a different distance function, e.g.

$$\begin{aligned} d_P\left(\boldsymbol{x}, \boldsymbol{x}^{\mathrm{s}}\right) &= \|P^{-1} \cdot \left(\boldsymbol{x} - \boldsymbol{x}^{\mathrm{s}}\right)\| \\ P &= \begin{pmatrix} 1 & 1 \\ 4\sqrt{a}/\lambda_+ & 4\sqrt{a}/\lambda_- \end{pmatrix}, \end{aligned} \tag{3.4}$$

with $\lambda_{\pm} = -b/2 \pm \sqrt{b^2/4 - 2\sqrt{a}}$ being the complex eigenvalues of the linearization of Equation (3.1). This distance function is related to the $\|\cdot\|_P$-norm in Proposition A.4.

For more complex attractors, e.g. the later treated chaotic Rössler attractor, this is not so straightforward unfortunately. However, we will present a pragmatic solution in Section 3.3.2. Hence, we developed Condition III.

Finally, $T_F^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ and $T_L^{\epsilon}(\boldsymbol{x}^{\mathrm{a}})$ do change under a smooth transformation of coordinates, because the Euclidean metric is not invariant. This implies that the result depends on the coordinates chosen for the system and might not be system-inherent. Hence, we demand Condition IV.

3.3 Two Complementary Metrics

To tackle the aforementioned problems and fulfill Conditions I to IV, we develop two metrics for general system, as in Equation (2.6): *Area under Distance Curve* (D) and *Regularized Reaching Time* (T_{RR}). They naturally lead to an analysis of the transient from two different points of view, as explained in the following.

3.3.1 Area under Distance Curve

Area under Distance Curve (D) derives from the idea that a trajectory stays distant from the attractor during the transient and close to it in the asymptotic phase. To have these notions of "distant" and "close", a distance function $d(\cdot, \cdot)$ is needed, and we define

$$D(x) = \int_0^\infty \mathrm{d}t \; d(\varphi(t, x), \mathcal{A}) \qquad \forall x \in \mathcal{X}. \tag{3.5}$$

This is the cumulative distance to the attractor. The influence of the asymptotics is removed due to the trajectory being close to the attractor for $t \to \infty$ for $x \in \mathcal{B}_\mathcal{A}$. If $x \in \mathcal{X} \backslash \mathcal{B}_\mathcal{A}$, then $D(x) = \infty$ because the corresponding trajectory $\phi(\cdot, x)$ never approaches $\mathcal{A}$.

Because Equation (3.5) is the integral over the distance in time, D is the area below the distance curve, hence its name. A different point of view is that D is the time weighted by the distance. The definition in Equation (3.5) using the limit of an integral demands for a note on convergence and that is given in Section 3.6.

As seen from the usage in Equation (3.5), the distance function $d(\cdot, \cdot)$ is between an element of state space $\mathcal{X}$ and the attractor $\mathcal{A}$. If $\mathcal{A}$ contains more than one point, it could be the infimum of the distances to all points within. By choosing a tailor-made function $d(\cdot, \cdot)$, one may adjust the metric to specific research questions, e.g. by letting $d(x, \mathcal{A})$ to be some sort of cost or damages due to being distant to the attractor $\mathcal{A}$. The main constraints on $d(\cdot, \cdot)$ are that it should be approaching 0 around the attractor and be 0 on it.

D can be estimated in a straightforward manner from the trajectory as the definition in Equation (3.5) was done using an integral. This is understood to be under the assumption that the attractor is known, which is a prerequisite for the chapter, so we can focus on the analysis of the transient.

$x \in \mathcal{X}$ with high values of D are called "reluctant" and ones with low values "eager". This terminology is used to emphasize the fact, that reluctant states go through a transient phase distant from the attractor while eager states approach it right away.

Straightforward differentiation gives the orbital derivative

$$\frac{\partial}{\partial t} D(\varphi(t, x)) = -d(\varphi(t, x), \mathcal{A}). \tag{3.6}$$

Hence, its value decreases strictly along the flow. This property will be important in Section 3.3.2.2. Furthermore, this implies D is a Lyapunov function [Gie2007]. Moreover, Equation (3.6) in combination with the additional condition $D(x) = 0 \; \forall x \in \mathcal{A}$ gives an alternative definition for D.

3.3.2 Regularized Reaching Time

The second novel metric to measure transient time is *Regularized Reaching Time* and it is founded on time differences between trajectories. Taken from the original article

[Kit2017f]: "It can be interpreted as the additional time (positive or negative) that a trajectory starting at a point of interest needs to approach the attractor after a reference trajectory has already approached it. A positive or negative value means that the trajectory at hand approaches the attractor by this much later or earlier, respectively, than the reference trajectory does."

In order to formalize this idea, we introduce an auxiliary metric $t^\epsilon(\boldsymbol{x})$ as the time a trajectory starting at an arbitrary initial condition $\boldsymbol{x}$ needs to enter an ϵ-neighborhood around the attractor for some small ϵ. Hence, taking some function $\Delta\colon \mathfrak{X} \to \mathbb{R}_{\geq 0}$ then we want

$$\epsilon = \Delta(\varphi(t^\epsilon(\boldsymbol{x}), \boldsymbol{x})), \tag{3.7}$$

to hold, where we demand Δ to be 0 on the attractor and $\Delta(\varphi(t, \boldsymbol{x}))$ to be strictly and continuously decreasing in t. Hence, Δ plays the role of a generalized distance function, measuring how far a point $\boldsymbol{x} \in \mathfrak{X}$ is away from the attractor. $\varphi(t^\epsilon(\boldsymbol{x}), \boldsymbol{x}))$ is the state after starting at $\boldsymbol{x}$ and evolving the system for the time $t^\epsilon(\boldsymbol{x})$. Therefore, Equation (3.7) gives an implicit definition of $t^\epsilon(\boldsymbol{x})$ as the time to reach an ϵ-neighborhood around the attractor, with respect to the generalized distance function Δ.

Since the actual reaching time of a trajectory to an attractor may be infinite, T_{RR} is formally described by the difference of $t^\epsilon(\boldsymbol{x})$ and $t^\epsilon(\boldsymbol{x}^{\text{ref}})$ in the limit $\epsilon \to 0$ for some reference initial condition $\boldsymbol{x}^{\text{ref}}$,

$$T_{RR}(\boldsymbol{x}; \boldsymbol{x}^{\text{ref}}) = \lim_{\epsilon \to 0} \left(t^\epsilon(x) - t^\epsilon(\boldsymbol{x}^{\text{ref}}) \right). \tag{3.8}$$

If the attractor is a hyperbolic fixed point, we prove in Appendix A under mild conditions that there exists a class of choices for Δ such that the limit in Equation (3.8) exists for all $\boldsymbol{x}$ in the attractor's basin except the strong stable manifold and the attractor itself. The manifold associated to all Lyapunov exponents except the weakest one is what we call the strong stable manifold. Additionally, if T_{RR} exists, it is unique, meaning independent of the choice of Δ.

3.3.2.1 Properties

In Appendix A, we prove that T_{RR} is a parametrization of the strong stable foliation. In particular, this implies that after the application of a diffeomorphism Φ of the state space $\mathfrak{X}$, the diffeomorphic image of the strong stable foliation will parametrize the level sets of T_{RR}. Hence, T_{RR} is invariant under such transformations, which are nothing else than changes of coordinates,

$$T_{RR}(\Phi(\boldsymbol{x}); \Phi(\boldsymbol{x}^{\text{ref}})) = T_{RR}(\boldsymbol{x}; \boldsymbol{x}^{\text{ref}}). \tag{3.9}$$

$\boldsymbol{x}^{\text{ref}}$ had to be transformed, too, obviously.

States with relatively low value of T_{RR} are called "early" and with high value "late". This is done in order to emphasize that T_{RR} actually represents the time a trajectory approaches the attractor earlier (oder later) than the reference trajectory.

The influence of $\boldsymbol{x}^{\text{ref}}$ is not structural, but just shifts T_{RR} by a constant (i.e. $\boldsymbol{x}$-independent) value. This can be seen by taking Equation (3.8) for two different reference points $\boldsymbol{x}^{\text{ref}}$ and $\boldsymbol{x}^{\text{ref}\prime}$ and subtract one from the other

$$T_{RR}\left(\boldsymbol{x};\boldsymbol{x}^{\text{ref}}\right) - T_{RR}\left(\boldsymbol{x};\boldsymbol{x}^{\text{ref}\prime}\right) = T_{RR}\left(\boldsymbol{x}^{\text{ref}\prime};\boldsymbol{x}^{\text{ref}}\right). \tag{3.10}$$

As the RHS is independent of $\boldsymbol{x}$ there is no structural impact on T_{RR} when choosing different $\boldsymbol{x}^{\text{ref}}$. Hence, when analyzing T_{RR} over distributions of initial conditions in state space, central moments, that are invariant under shifts, are sensible and used later on. For completeness, it should be mentioned that

$$T_{RR}\left(\boldsymbol{x}^{\text{ref}};\boldsymbol{x}^{\text{ref}}\right) = 0, \tag{3.11}$$

which is just a check of consistency.

A restriction for the choice of $\boldsymbol{x}^{\text{ref}}$ is that it had better not be chosen on the attractor. The reason is that if $\boldsymbol{x}^{\text{ref}} \in \mathcal{A}$ then $t^{\epsilon}(\boldsymbol{x}^{\text{ref}}) = 0\ \forall \epsilon > 0$ but for the points $\boldsymbol{x} \in \mathcal{X}\backslash\mathcal{A}$ outside the attractor $t^{\epsilon}(\boldsymbol{x})$ still goes to ∞ for $\epsilon \to 0$, implying that $T_{RR}\left(\boldsymbol{x};\boldsymbol{x}^{\text{ref}}\right) = -\infty$. This result is reasonable, because a trajectory starting on the attractor stays there while an outside trajectory may only approach it without ever reaching it. Similarly, one can argue why it should not lie on the strong stable manifold either.

To derive an orbital derivative, we can start with $T_{RR}(\varphi(t,\boldsymbol{y});\boldsymbol{x}^{\text{ref}})$ (with some initial condition $\boldsymbol{y} \in \mathcal{X}$) and take the derivative with respect to t the formal definition of T_{RR} in Equation (3.8), apply the formal definition Equation (3.8) and take the derivative with respect to t

$$\frac{\partial}{\partial t} T_{RR}(\varphi(t,\boldsymbol{y});\boldsymbol{x}^{\text{ref}}) = \lim_{\epsilon\to 0} \frac{\partial}{\partial t} t^{\epsilon}(\varphi(t,\boldsymbol{y})), \tag{3.12}$$

where we assumed the exchangeability of the limit and the derivative. Next, plugging in $\boldsymbol{x} = \varphi(t,\boldsymbol{y})$ in Equation (3.7) and taking the derivative w.r.t t yields (with an appropriate application of the chain rule and sorting)

$$0 = \left(\partial_1 \left(\Delta \circ \varphi\right)\right)\left(t^{\epsilon}(\varphi(t,\boldsymbol{y})) + t, \boldsymbol{y}\right) \cdot \left(\frac{\partial}{\partial t} t^{\epsilon}(\varphi(t,\boldsymbol{y})) + 1\right). \tag{3.13}$$

By definition, $\Delta(\varphi(t,\boldsymbol{x}))$ is strictly decreasing in t, hence its derivative is negative

$$\frac{\partial}{\partial t}\Delta(\varphi(t,\boldsymbol{x})) = \left(\partial_1 \left(\Delta \circ \varphi\right)\right)(t,\boldsymbol{x}) < 0 \tag{3.14}$$

and in particular non-zero. Thus, combining Equations (3.12) and (3.13) yields the orbital derivative of T_{RR}

$$\frac{\partial}{\partial t} T_{RR}(\varphi(t,y);\boldsymbol{x}^{\text{ref}}) = -1. \tag{3.15}$$

While it looks trivial at first, it is actually the natural outcome: T_{RR} actually represents a time to approach the attractor, so the change in values along a trajectory should be

precisely the time passed. Equation (3.15) demonstrates that T_{RR} is a Lyapunov function [Gie2007].

To use Equation (3.15) as an alternative definition we need another constraint. Because of $T_{RR}(\boldsymbol{x}; \boldsymbol{x}^{\text{ref}}) = -\infty \ \forall \boldsymbol{x} \in \mathcal{A}$, this cannot be done on the attractor (in contrast to D). In case of hyperbolic fixed points, it follows directly from Proposition A.7 that T_{RR} is a parametrization of the strong stable foliation $\mathcal{F}_{ss}$, whose definition is recalled in Theorem A.6. So we can use the constraint that $T_{RR}(\boldsymbol{x}; \boldsymbol{x}^{\text{ref}}) = 0 \ \forall \boldsymbol{x} \in \mathcal{F}_{ss}$, where we call $\mathcal{F}_{ss}^{ref} = \mathcal{F}_{ss}(\boldsymbol{x}^{\text{ref}})$ the *reference leaf* containing $\boldsymbol{x}^{\text{ref}}$. For more complex attractors, a generalized condition needs to be found and this is part of the outlook.

Proposition A.4 provides the convergence of T_{RR} in Equation (3.8) for hyperbolic fixed points only. When thinking about more complex attractors that may arise in real-world examples the question of convergence comes up again. A general idea why T_{RR} should converge with a well chosen Δ in this case, too, is that in the asymptotics, trajectories will "behave similarly" because they are close to the attractor. So, for two very small $\epsilon_1 > \epsilon_2$, the time difference to enter the ϵ_2-environment after entering the one of ϵ_1 should be roughly the same, independent from where a trajectory started. Hence, for two states $\boldsymbol{x}$ and $\boldsymbol{x}^{\text{ref}}$ we can assume $t^{\epsilon_2}(x) - t^{\epsilon_1}(x) \approx t^{\epsilon_2}(\boldsymbol{x}^{\text{ref}}) - t^{\epsilon_1}(\boldsymbol{x}^{\text{ref}})$ implying $t^{\epsilon_2}(x) - t^{\epsilon_2}(\boldsymbol{x}^{\text{ref}}) \approx t^{\epsilon_1}(x) - t^{\epsilon_1}(\boldsymbol{x}^{\text{ref}})$. This suggests that the limit in Equation (3.8) might exist. So a crucial problem is to find an appropriate function for Δ in order to get an estimation for T_{RR}.

3.3.2.2 Estimation

The first idea as a choice for a Δ would be the infimum of the Euclidean distance to the points in the attractor. Basically, this means that t^{ϵ} should be replaced by T_F^{ϵ} or T_L^{ϵ} from Section 3.2.1. This would give a very coarse estimation but is probably not the correct choice as the condition of Δ being strictly decreasing along the flow is in general not fulfilled.

A pragmatic choice of Δ is D, the Area under Distance Curve. It fulfills both conditions demanded for Δ (cf. Section 3.3.1) when using for $d(\cdot, \cdot)$ the infimum of the Euclidean distance to the attractor points. Hence, for ϵ small enough, we can define $t_D^{\epsilon}(\boldsymbol{x})$ as the time until the D (Equation (3.5)) of the trajectory's remainder is ϵ-small

$$\epsilon = D\left(\varphi(t_D^{\epsilon}(\boldsymbol{x}), \boldsymbol{x})\right) = \int_{t_D^{\epsilon}(\boldsymbol{x})}^{\infty} \mathrm{d}t \ d\left(\varphi(t, \boldsymbol{x}), \mathcal{A}\right). \tag{3.16}$$

Note that the ideas for D and T_{RR} are generally independent and the usage of D in this case is purely because it fulfills the above mentioned conditions. So it is a good, pragmatic choice.

Using t_D^{ϵ} defined in Equation (3.16) as the time-function t^{ϵ} in Equation (3.8) for the estimation of T_{RR}, our numerical results show that this idea is sensible for more complex attractors, e.g. in the Rössler system below.

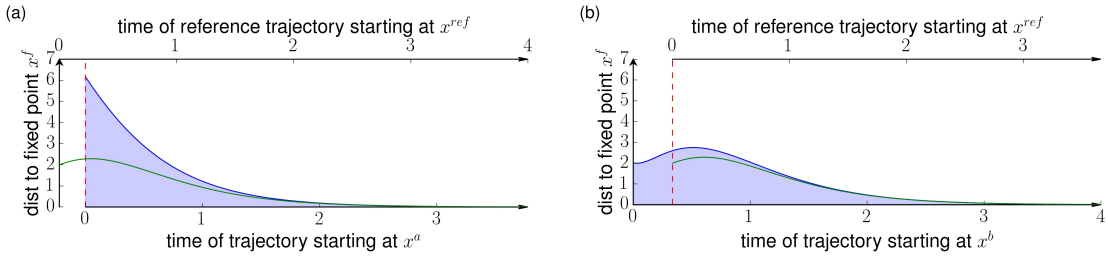


Figure 3.2: The figure shows for two exemplary initial conditions (a) $x^{\mathrm{b}} = (0.8, 2.35)^\top$ and (b) $x^{\mathrm{c}} = (1.4, 0.24)^\top$ the distance of the attractor over time (blue curve) in the linear example system of Section 3.4.1. The initial conditions have been chosen such that the D value, which corresponds to the blue-shaded area, is the same for both trajectories, $D(x^{\mathrm{b}}) = D(x^{\mathrm{c}}) = 3.8$. But the trajectory starting at x^{b} approaches the attractor *earlier* than the reference trajectory (green in (a) and (b)), which in turn is *earlier* than the one from x^{c}, meaning $T_{RR}(x^{\mathrm{b}}; x^{\mathrm{ref}}) = -0.22 < T_{RR}(x^{\mathrm{ref}}; x^{\mathrm{ref}}) = 0 < T_{RR}(x^{\mathrm{c}}; x^{\mathrm{ref}}) = +0.34$. In order to show this, the example trajectories (blue) have been shifted in each plot by the value of T_{RR} with respect to the reference trajectory (green). This demonstrates an intuition behind T_{RR}: it describes by how much one has to shift one trajectory so it matches the asymptotics of the reference trajectory. These graphics were published in [Kit2017f].

3.4 Examples

In order to demonstrate the applicability of the metrics, we selected four examples with differing properties and increasing complexity.

3.4.1 Linear System

Even though we want to focus on going in the direction of application to real-world systems, understanding some features in a basic linear system proves useful. For general systems, T_{RR} and D can be tackled numerically only. But a linear system can be solved analytically and explicit expressions for both metrics were found. We will first analyze both metrics for a general linear system and then discuss a chosen example.

T_{RR} for a general linear system. For a hyperbolically stable linear system with a (complex-) diagonalizable matrix $A \in \mathbb{R}^{n \times n}$ and the fixed point x^{f} at the origin,

$$\dot{x} = A \cdot x, \tag{3.17}$$

we decompose $x = \sum_{i=0}^{n-1} \alpha^i v^{\mathrm{i}}$ with coefficients $\alpha^0, \ldots, \alpha^{n-1}$ in the eigenvector basis $v^0, \ldots, v^{\mathrm{n-1}}$ with eigenvalues $\lambda^0, \ldots, \lambda^{n-1}$ sorted in descending order by real part. We assume in particular λ^0 to have a strictly larger real part than λ^1 and multiplicity one. Hence we can apply Equation (A.14) derived in Appendix A and get

$$T_{RR}(x; x^{\mathrm{ref}}) = \frac{1}{\lambda_0} \ln \left| \frac{\alpha^{0,\mathrm{ref}}}{\alpha^0} \right|, \tag{3.18}$$

where $\alpha^{0,ref}$ is the α^0 coefficient for the reference point x^{ref}. $\alpha^{0,ref}$ should be non-zero, i.e. x^{ref} should not be on the strong stable manifold.

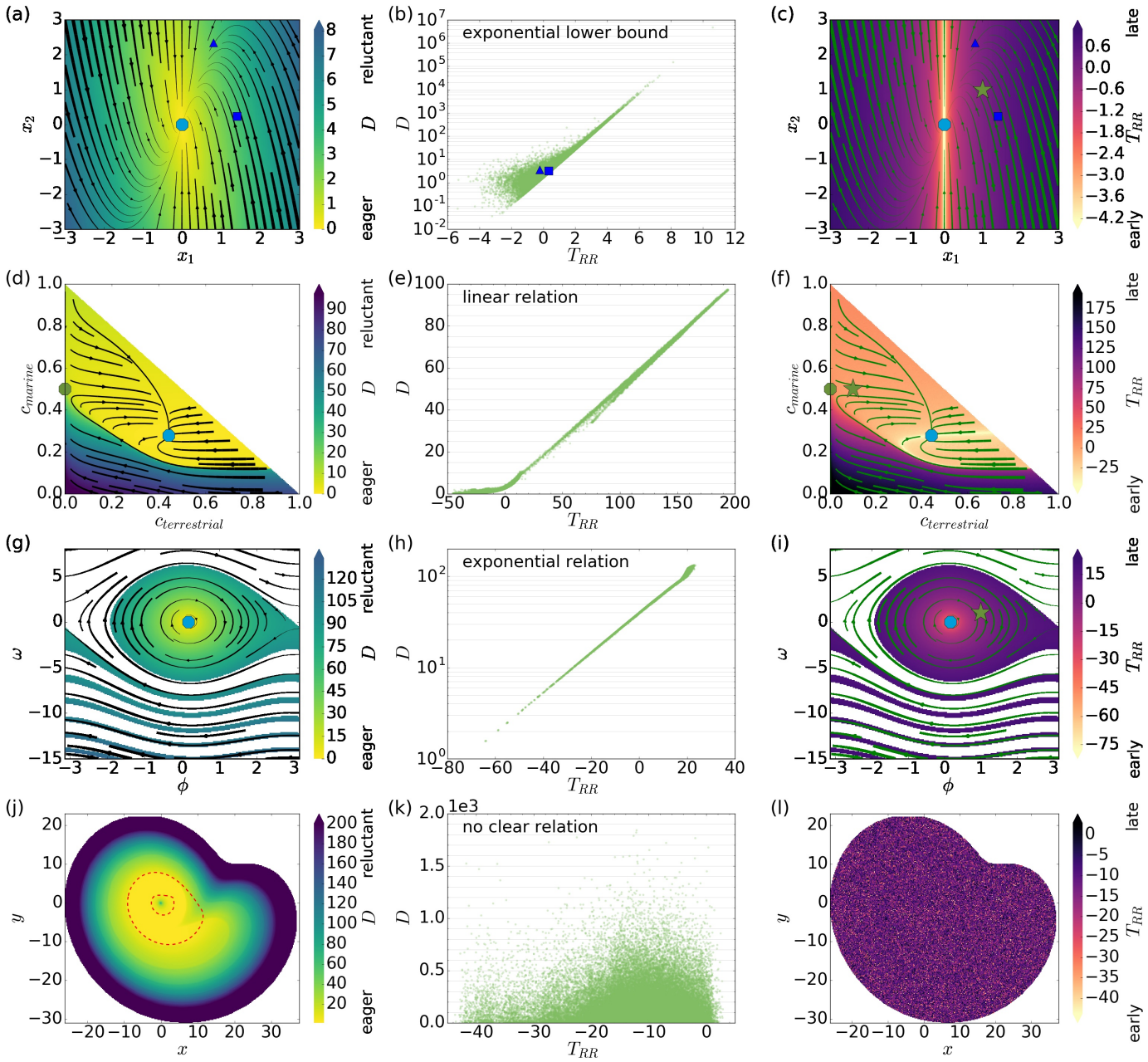


Figure 3.3: For the presented example systems (top to bottom: linear system, global carbon cycle, generator in a power grid, Rössler system) the two new metrics have been computed for each initial condition in the state space and marked with color, see left column *Area under Distance Curve* (D) and right column *Regularized Reaching Time* (T_{RR}). The middle column shows their relations for the particular system. The initial conditions x^{b} (triangle) and x^{c} (square) from Figure 3.2 have been marked in (a-c), too. As the Rössler system is 3-dimensional, the above plot depicts only a slice at fixed $z = 0.6$ where the boundary of the attractor's projection to this plane is shown in dashed red lines. As this is only a projection, D is not 0 for all points within. For comparison, graphical representations of the full attractor can be found in Figures 2.1 and 2.2. These graphics were published in [Kit2017f].

Note that Proposition A.4 gives the uniqueness of this result independent of the choice of Δ.

In Equation (3.18), T_{RR} depends only on α^0, meaning the projection of x on the eigenvector corresponding to the least stable eigenvalue λ^0. While this might be counter-intuitive in the beginning, it can be explained: The contributions from all other eigenvalues are vanishing because they decay faster than λ^0 by definition. So for a linear system, only the contribution from λ^0 remains. Also, on the strong stable manifold where $\alpha^0 = 0$, the values for T_{RR} go to $-\infty$ which we mentioned already in Section 3.3.2.1 for general systems.

D for a general linear system. Taking the system (3.17) and choosing $d\left(x, \{x^{\mathrm{f}}\}\right) = d_E\left(x, x^{\mathrm{f}}\right)^2$ the squared Euclidean distance, we calculate D directly by using the definition Equation (3.5)

$$D(x) = \sum_{i,j=0}^{n-1} \frac{-(\alpha^i)^* \alpha^j}{(\lambda^i)^* + \lambda^j} (v^{\mathrm{i}})^\dagger v^{\mathrm{j}}. \tag{3.19}$$

Therefore, in case of D, all eigenvalues contribute, contrary to T_{RR}. But they are weighted as can be seen in the denominator. In case of A being symmetric, this formula can be reduced to $D(x) = \frac{1}{2} x^\top A^{-1} x$.

T_{RR} for an example linear system. We choose the $n = 2$-dimensional linear system

$$\dot{x} = \begin{pmatrix} -1 & 0 \\ 4 & -2 \end{pmatrix} \cdot x \tag{3.20}$$

with a stable and a strong stable eigenvalue and corresponding eigenvectors

$$\lambda^s = -1,\ v^{\mathrm{s}} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \text{ and } \lambda^{ss} = -2,\ v^{\mathrm{ss}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.21}$$

We choose the reference point to be $x^{\mathrm{ref}} = (1, 1)^\top$. Identifying $\lambda^0 = \lambda^s$, $v^0 = v^{\mathrm{s}}$ implies $\alpha^0 = x_0$. Then, using Equation (3.18) gives

$$T_{RR}(x; x^{\mathrm{ref}}) = \ln\left(|x_0|\right). \tag{3.22}$$

This result is also visible in the numerical estimation in Figure 3.3 (c); the values of T_{RR} change only in the direction of x_0. The coloring describes the values of the metrics (cf. the colorbar in the right of the figures) and the green star represents x^{ref}.

In order to get a better feeling for these metrics, we have chosen two exemplary initial conditions, an *early-eager* one and a *late-eager* one, and plotted their trajectories' distance to the attractor over time in Figure 3.2. We see an intuition for T_{RR}: it can be interpreted as the time-shift between the original trajectory and the reference trajectory until the asymptotics match. So we plotted both trajectories shifted to each other using the analytical result for T_{RR} in Equation (3.22).

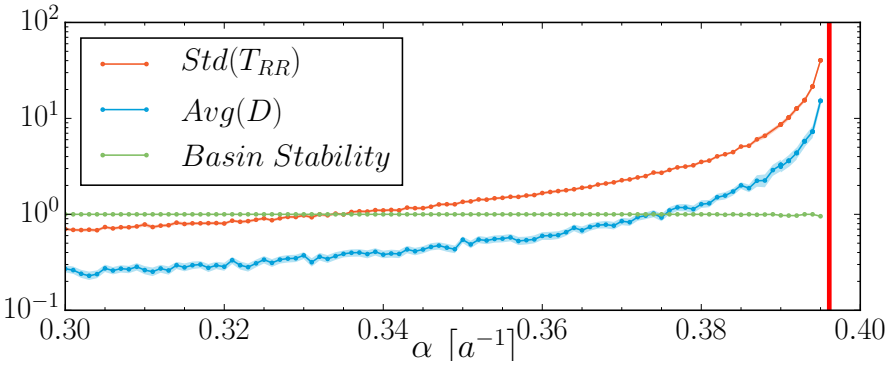


Figure 3.4: For the *global carbon cycle* in Equations (3.25a) and (3.25b), the mean of D and standard deviation of T_{RR} are plotted (with their 5%-95% bootstrap error) and show a divergence before the parameter α (yearly human carbon offtake) reaches the bifurcation value (marked by the red line). For comparison Basin Stability is shown, which does *not* show any change because the size of the basin stays constant before the bifurcation. This graphic was published in [Kit2017f].

D for an example linear system. Analyzing D for the example linear system in Equation (3.20) gives

$$D(x) = \frac{11}{6}x_0^2 + \frac{1}{4}x_1^2 + \frac{2}{3}x_0x_1, \tag{3.23}$$

where Equation (3.19) has been used. The numerical result in Figure 3.3 (a) confirms this.

In Figure 3.2, the blue-shaded area corresponds to the D value which is the same in both cases of our particular choice. This choice was made in order to see how trajectories can have differing T_{RR} values even if the D values match.

The exponential lower bound that comes up in the scatter plot Figure 3.3 (b) can be calculated analytically by combining Equations (3.22) and (3.23)

$$D(x) \geq \frac{25}{18} e^{2T_{RR}(x;x^{\text{ref}})}. \tag{3.24}$$

3.4.2 Global Carbon Cycle

The second example has been chosen to take a step in the direction of real-world examples. It is a conceptual model of the global carbon cycle proposed by Anderies et al. [And2013]. The details are in Model Description 2. For the analysis we refer to the flow (streams) that is drawn in Figures 3.3 (d) and 3.3 (f) and the α parameter stated above. The whole phase space of Equations (3.25a) and (3.25b) is the basin of attraction of the fixed point in the middle marked by a blue dot. The trajectories starting in the lower part have to pass by a "desert-like" saddle (with $c_t = 0$) at the left (green dot).

Model Description 2 (Global Carbon Cycle)**.** In [And2013] a low-dimensional model of the global carbon cycle was developed and we use the pre-industrialization version. It consists of three dynamical variables, the terrestrial, marine and atmospheric carbon stocks, denoted by $c_t = c_{terrestrial}$, $c_m = c_{marine}$ and $c_a = c_{atmospheric}$ respectively. Furthermore, the conservation of total carbon is formulated in the constraint $C = c_t + c_m + c_a = const.$ Thus, we can reduce the system to 2 state variables c_t and c_m and rescale the units such that $C = 1$:

$$\dot{c}_t = NEP(p, r, c_t) - \alpha c_t \tag{3.25a}$$
$$\dot{c}_m = I(c_a, c_m) \tag{3.25b}$$

$$NEP(p, r, c_t) = r_{tc} c_t \cdot \left(1 - \frac{c_t}{k}\right)(p - r) \tag{3.25c}$$
$$p = g a_p T^{b_p} e^{-c_p T} \tag{3.25d}$$
$$r = a_r T^{b_r} e^{-c_p T} \tag{3.25e}$$
$$T = a_T c_a + b_T \tag{3.25f}$$
$$g = a_f {c_a}_f^b \tag{3.25g}$$
$$I(c_a, c_m) = a_m \cdot (c_a - \beta c_m), \tag{3.25h}$$

where NEP is the net Eco-system production, p photosynthesis, r respiration, α harvesting parameter and I diffusion; indirect dependencies have been omitted and more details are in [And2013; Hec2016]. Implicit dependencies are only given for the main differential equations in order to keep the equations understandable. The parameters, as estimated in [And2013], are summarized in Table 3.1. More details are in the original work [And2013] and a detailed analysis including a slight extension can be found in [Hec2016].

The color in Figure 3.3 (f) depicts T_{RR} and the first finding is the splitting of the basin of attraction. The strong stable manifold of the stable node becomes visible as a light beige line due to its low values of T_{RR}, i.e. as very *early* states because $T_{RR} \to -\infty$. So it is the separatrix for the observed splitting. Also, the expected smooth increase of the return times when distancing (along the trajectories) from the attractor can be observed.

Still, the splitting of the basin of attraction is visible for values of $c_{terrestrial} < 0.3$, where it is only due to quantitatively different behavior and the visible boundary is actually a rather sharp but still continuous transition. (The latter statement follows right from Theorem A.6 and Proposition A.7). Looking at Figure 3.3 (f) one can also see that the boundary becomes more and more fuzzy for even smaller values of $c_{terrestrial}$, demonstrating that there is really a need for a quantitative analysis.

When applying D to this model (Figure 3.3 (d)), the splitting of the basin can be observed again. In contrast to T_{RR}, the strong stable manifold of the stable node is not

Parameter	Value	Parameter	Value
a_f	1.5	a_T	0.8
b_f	0.3	b_T	0.2
a_r	110	t_{tc}	2.5
a_p	220	k	0.7
b_r	4	a_m	0.05
b_p	3	β	1
c_r	5	α	0.3^*
c_p	7		

Table 3.1: Set of Parameters for the global carbon cycle estimated in [And2013]. The value of α is marked with a * because it is varied for the bifurcation analysis in Figure 3.4 and the TSM-analysis in Section 4.4. This table was published in the supplemental data of [Kit2017f] based on [And2013].

visible because D can be seen as a (by distance) weighted time and the contributions from the asymptotic part where the difference in the Lyapunov spectrum matters are negligible.

Furthermore, we see a clear linear correlation of both metrics in Figure 3.3 (e) because all trajectories starting in the lower part have to pass by at the saddle on the left and spend a long time there.

Both metrics work as *early-warning signals* [Len2011; Sch2009], too. When increasing α, corresponding to the harvest of terrestrial carbon, the system passes through a subcritical pitchfork bifurcation where the saddle becomes stable and the lower-left part of the phase space splits off. The divergences of the two metrics' statistics as seen in Figure 3.4 prove their prebifurcational sensitivity, while other systemic indicators like basin stability [Men2013] do not change (up to numerical fluctuations, see Figure 3.4). Note that in this example, a Lyapunov exponent analysis of the saddle would be able to predict the bifurcation due to the simplicity of the saddle also. However, in case of a more complex saddle, this would become arbitrarily difficult while this numerical estimation would still be possible for both metrics.

3.4.3 Generator in a Power Grid

As the next example, we chose the swing equation in Equation (3.26), a basic model describing the dynamics of a single generator connected to a large power grid, see Model Description 3. The parameters for the example have been chosen as follows: net power production $P = 1$ (at the node), capacity of the transmission line $K = 6$ and dampening $\alpha = 0.1$.

Model Description 3 (Single Generator in a Power Grid)**.** The Swing equation is a minimalistic model describing a single generator within a power grid, also being used in electrical engineering [Wec2013; Yua2003]. It consists of two dynamical variables, the phase θ and angular frequency ω, both in a reference frame rotating at the grid's rated frequency. The parameters of the system correspond to the net power production P (at the node), the capacity of the transmission line K and dampening α.

$$\begin{aligned} \dot{\phi} &= \omega, \\ \dot{\omega} &= 2P - \alpha\omega - 2K \sin\phi. \end{aligned} \tag{3.26}$$

It is formally equivalent to a pendulum with constant driving and damping. The stable fixed point at $\omega^s = 0$, $\phi^s = \arcsin \frac{P}{K}$, if it exists, describes a state of synchronization. Depending on the choice of the parameters, the system may exhibit another attractor: a limit cycle at larger positive values of ω. For negative values, the two basins of attraction are then interleaved. A more detailed introduction and analysis can be found in [Men2014; Sch2014; Wec2013].

Calculating T_{RR} inside the basin of the stable fixed pointed (ω^s, θ^s) yields Figure 3.3 (i). There is basically no color change away from the attractor, so we can see that a trajectory barely spends any time in the transient and goes quickly to the attractor. Analogously, Figure 3.3 (g) for D leads to the same conclusion as T_{RR}.

Comparing both metrics in Figure 3.3 (h) shows that they are closely linked. Note that this time D is presented on a logarithmic scale, so the relation is exponential and what we see here is actually the influence of the linearized part of the system. The accumulation in the upper right corresponds to the initial conditions with lower values of ω. This means, they only go through a very short transient and spend most of their time in the part where the linearization holds.

The white parts in the phase spaces Figures 3.3 (g) and 3.3 (i) correspond to the basin of attraction of the limit cycle. As this means the system is away from synchrony, the generators would usually switch off before reaching it. So we did not include it in the analysis.

3.4.4 Chaotic Rössler Oscillator

Although we have proven the convergence of T_{RR} for fixed points only, we show with the chaotic Rössler system that both metrics are applicable to higher-dimensional and more complex attractors, too. Details of the model are in Model Description 1 on page 9.

Figure 3.3 (l) shows a slice of the phase space with the standard parameters $a = 0.2$, $b = 0.2$, $c = 5.7$ for T_{RR} and the expected sensitivity to initial conditions for chaos is observed: *early* and *late* trajectories lie closely together and the metric T_{RR} has low spatial correlation.

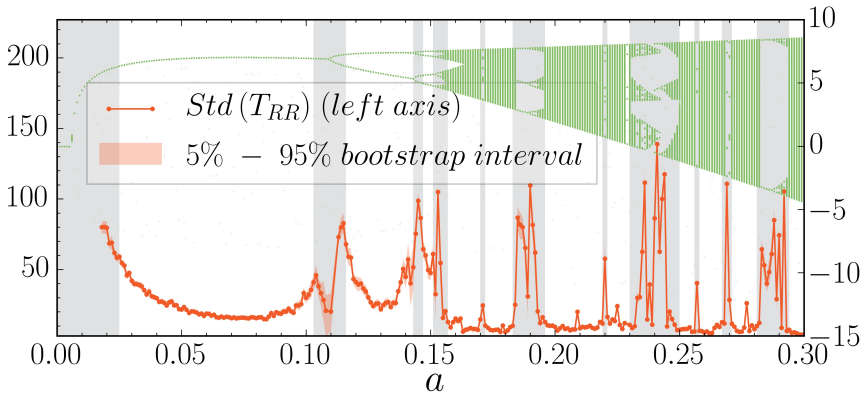


Figure 3.5: The bifurcation diagram (green) of the Rössler system for varying the parameter a in Equation (2.14) was computed from the local maxima in z of the attractor and T_{RR} (orange) shows a strong sensitivity to these qualitative changes. The gray background is used so the reader can more easily connect the peaks in T_{RR} to the corresponding parts in the bifurcation diagram. This graphic was published in [Kit2017f].

In contrast, D shows in Figure 3.3 (j) surprisingly smooth changes of an embryo-like shape. Because the focus of this article is on transient dynamics a new feature of the chaotic Rössler system is uncovered: while the attractor is chaotic, the basin of attraction is very regular. D focuses on the initial transient and the chaotic asymptotics is filtered out. For comparison, the boundaries of the attractor's projection have been added with dashed red lines in Figure 3.3 (j) and depictions of the attractor are in Figures 2.1 and 2.2.

Furthermore, T_{RR} can be applied as an early-warning signal in this case, too. In order to demonstrate this, we chose to vary a as it has a crucial influence on the system's dynamics (see the bifurcation diagram in Figure 3.5 (green)). For values of $a < 0.006$ (cf. [Bar2011]) there is only a single stable fixed point. At $a \approx 0.006$ a limit cycle emerges due to a Hopf bifurcation [Bar2011]. For $a > 0.11$, several period doublings are observed, finally leading to chaos for $a > 0.155$. Even in the chaotic regime, further bifurcations can be observed.

In Figure 3.5, the standard deviation of the T_{RR} distribution from randomly chosen initial conditions inside the basin of attraction is given. Due to the sensitive dependence on initial conditions, the reference value varies a lot and hence introduce shifts in the distribution that do not describe actual changes in the system's dynamics. To remove this effect, it is crucial to use (functions of) central moments like the standard deviation.

T_{RR} is strongly sensitive to any qualitative changes in the dynamics of the system, incl. even chaos-chaos transitions. Closely observing Figure 3.5 uncovers that there is a base-line with little fluctuations at $Std(T_{RR}) \approx 10$ complemented with strong peaks. In the chaotic regime, the peaks correspond directly to qualitative changes. Also, we observe sensible changes during the period-doubling phase and a strong increase before the Hopf bifurcation at $a \approx 0.006$, proving the usefulness as an *early-warning signal*.

3.5 Timing on Networks

Remark. The following section refers to dynamical systems with an underlying network structure. Hence, as specified in Section 2.2.1, the general equation governing this dynamics is Equation (2.9), we assume the existence of an adjacency matrix A and the index of a vector $\boldsymbol{x} \in \mathfrak{X}$ refers to the smaller vector $\boldsymbol{x}_i \in \mathfrak{X}_i$ associated to the i^{th}-node. So assuming they have the same intrinsic dynamics, the full dimension of the system $n = \dim \mathfrak{X}$ is the product of the dimension of the dynamics on a single node $\nu = \dim \mathfrak{X}_i$ and the number of nodes N. The set of nodes is $\mathcal{N}$.

In this section, we extend the idea of T_{RR} to networks. The idea is as follows: We have the metric T_{RR} (that depends on the state) and take the average over the subspace corresponding to a certain node i of the network system, while keeping the coordinates for the other nodes at the values for their attractor. That way, we associate a value to each node, creating a network metric (cmp. Section 3.1).

To put that in equations, we take the (normalized) invariant density μ (cf. Section 2.2.2) of the attractor of interest $\mathcal{A}$ and an assumed distribution of perturbations $\rho_i : \mathfrak{X}_i \to \mathbb{R}_{\geq 0}$ for each node i and find the *distribution of initial conditions* ξ that we want to average over

$$\begin{aligned} \xi_i(\boldsymbol{x}) &= \int_{\mathfrak{X}_i} \mathrm{d}\boldsymbol{y}_i \; \rho_i(\boldsymbol{x}_i - \boldsymbol{y}_i)\, \mu(\boldsymbol{y}) \\ \text{with } \boldsymbol{y} &= (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \boldsymbol{y}_i, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n)\,. \end{aligned} \tag{3.27}$$

The equation states that $\xi_i(\boldsymbol{x})$ is the integral over all combinations of states on the attractor $\boldsymbol{y}$ plus a perturbation $\boldsymbol{z}$, that lead to the position $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$. As we want to have perturbations only on one node i, ξ the integral is only over the i-th component. Note that in case of a fixed-point $\boldsymbol{x}^{\mathrm{f}}$, the invariant density would be an n-dimensional Dirac delta function, so the corresponding distribution of initial conditions ξ_i^f is simply the distribution of perturbations on that node (shifted appropriately to be centered around $\boldsymbol{x}^{\mathrm{f}}$) times an $(n-\nu)$-dimensional Dirac delta function for the other nodes,

$$\xi_i^f(\boldsymbol{x}) = \rho_i\left(\boldsymbol{x}_i - \boldsymbol{x}_i^{\mathrm{f}}\right) \prod_{j \in \mathcal{N},\; j \neq i} \delta\left(\boldsymbol{x}_j - \boldsymbol{x}_j^{\mathrm{f}}\right). \tag{3.28}$$

Having defined ξ in Equation (3.27) enables us to state T_{RR} as a network metric

$$T_{RR}^{\mathcal{N}}(i) = \frac{\int\limits_{\mathcal{B}_{\mathcal{A}}} \mathrm{d}\boldsymbol{x} \; \xi_i(\boldsymbol{x})\; T_{RR}\left(\boldsymbol{x}; \boldsymbol{x}^{\mathrm{ref}}\right)}{\int\limits_{\mathcal{B}_{\mathcal{A}}} \mathrm{d}\boldsymbol{x} \; \xi_i(\boldsymbol{x})}. \tag{3.29}$$

In [Mit2017b], we used the symbol $\langle T_R^1(i) \rangle$ instead of $T_{RR}^{\mathcal{N}}(i)$ and coined it *single-node recovery time*. I decided to adjust the symbol to match the rest of the dissertation. The term "recovery" instead of "reaching" is used to emphasize that this metric comes from the perspective on perturbations to a running system. In contrast, T_{RR}, as a metric that

depends on the initial condition, can be interpreted as both, the recovery time or the reaching time. Also, instead of using the algorithmic description in [Mit2017b], I decided to find a new mathematical description and the result is Equation (3.29).

3.5.1 Fixing the Reference Point

Extending T_{RR} to networks opens the question about the reference point $\boldsymbol{x}^{\mathrm{ref}}$ again. Now, this can be solved in an elegant manner by revisiting a remark in Section 3.3.2.1: a different choice of reference point leads to a constant shift in T_{RR} but does not change the function's structure, see Equation (3.10). But this statement can also be inverted: Adding a constant τ to T_{RR} is implicitly the same as choosing a different reference point. This can be seen by taking $\boldsymbol{x}^{\mathrm{ref}'} = \varphi\left(\tau, \boldsymbol{x}^{\mathrm{ref}}\right)$, then follows straight from the definition of T_{RR} that $T_{RR}\left(\boldsymbol{x}^{\mathrm{ref}'}; \boldsymbol{x}^{\mathrm{ref}}\right) = -\tau$ and plugging this in Equation (3.10) gives

$$T_{RR}\left(\boldsymbol{x}; \boldsymbol{x}^{\mathrm{ref}'}\right) = T_{RR}\left(\boldsymbol{x}; \boldsymbol{x}^{\mathrm{ref}}\right) + \tau. \tag{3.30}$$

Furthermore, this result implies that not only T_{RR} can be shifted by a constant value, but the same holds for $T_{RR}^{\mathcal{N}}$. So instead of choosing $\boldsymbol{x}^{\mathrm{ref}}$ explicitly, we can demand

$$T_{RR}^{\mathcal{N}}\left(i_{\min}\right) = 0 \qquad \text{with } i_{\min} = \operatorname*{argmin}_{i} T_{RR}^{\mathcal{N}}\left(i\right). \tag{3.31}$$

This means, the lowest value of $T_{RR}^{\mathcal{N}}$ for the node $i_{\min}$ is set to be zero and all the other nodes $i \neq i_{\min}$ have non-negative values. These describe how much *later* a trajectory approaches the attractor after a perturbation at this node (on average) as compared to perturbations at node $i_{\min}$.

3.5.2 Estimation

In order to estimate $T_{RR}^{\mathcal{N}}$ for a network with nodes $\mathcal{N}$, we can follow the algorithmic description proposed in [Mit2017b]:

Step 1 For each node $i \in \mathcal{N}$:

- **Step 1.1** We choose a random set of initial conditions $\mathcal{I}$ and for each $\boldsymbol{x} \in \mathcal{I}$ we go through the following steps:
  - **Step 1.1.1** Numerically integrate the trajectory from there for a long time, ensuring the final point $\boldsymbol{y}$ is very close to the attractor.
  - **Step 1.1.2** Draw a perturbation $\boldsymbol{\varkappa}$ from the distribution of perturbations ρ_i to find the new initial condition $\boldsymbol{y} + \boldsymbol{\varkappa}$.
  - **Step 1.1.3** If $\boldsymbol{y} + \boldsymbol{\varkappa}$ is inside $\mathscr{B}_{\mathscr{A}}$, then numerically integrate from $\boldsymbol{y} + \boldsymbol{\varkappa}$, until the trajectory is very close to the attractor. In this case, very close means for a fixed small ϵ and Δ as described in Section 3.3.2. If not, restart at step 1.1.1 with the next element in $\mathcal{I}$.
- **Step 1.2** We take the average of all the integration times in step 1.1.3, yielding a *pre-estimate* $T_{RR}^{\mathcal{N},pre}\left(i\right)$.

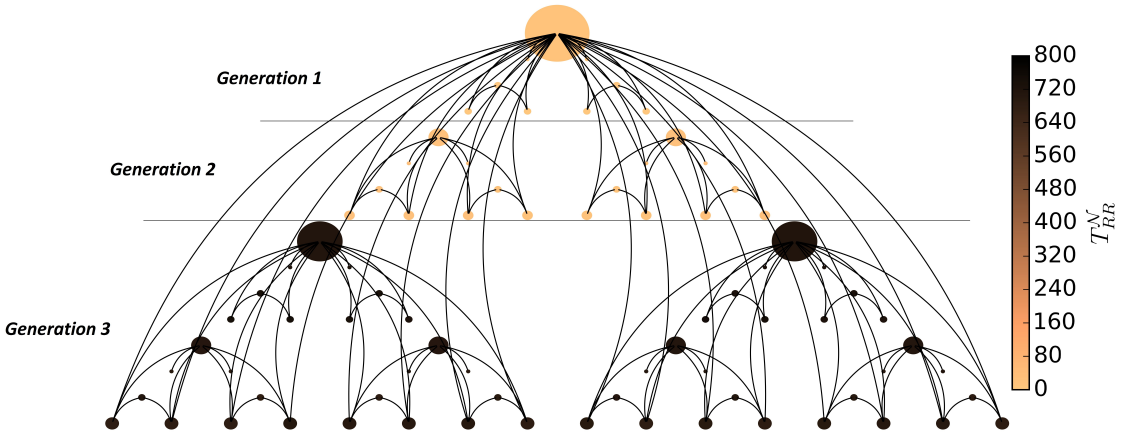


Figure 3.6: The network topology of the undirected deterministic scale-free network over three generations, i.e. with $N = 81$ identical Rössler oscillators. The size of each node is proportional to its degree and the color indicates the $T_{RR}^{\mathcal{N}}$ value of the corresponding node. This graphic was published in [Mit2017b].

Step 2 Finally, we subtract $\min_{j\in\mathcal{N}} T_{RR}^{\mathcal{N},pre}(j)$, finding for all $i \in \mathcal{N}$ the estimate

$$T_{RR}^{\mathcal{N}}(i) = T_{RR}^{\mathcal{N},pre}(i) - \min_{j\in\mathcal{N}} T_{RR}^{\mathcal{N},pre}(j)\,. \tag{3.32}$$

The steps in this algorithm can be related to the formal definitions before. Step 1.1.1 represents (when done for many points) the approximation of the invariant density of the attractor. Step 1.1.2 refers to the approximation of ξ_i in Equation (3.27). Step 1.1.3 computes the reaching time. Then, step 1.2 represents the averaging weighted with the distribution ξ_i as described in Equation (3.29). And finally, step 2 implements the condition Equation (3.31).

3.5.3 Examples

3.5.3.1 Deterministic scale-free network of Rössler oscillators

As the first example for *Timing of Networks*, we consider coupled Rössler Oscillators on a deterministic scale-free network [Bar2001] over three generations, hence in total $N = 81$ nodes. The structure of the network is depicted in Figure 3.6 and details of the dynamical model are in Model Description 4.

Model Description 4 (Network of coupled Rössler Oscillators)**.** In this model, multiple Rössler Oscillators, as described in Model Description 1, are linearly coupled on a network via their y component. Please note that as we are building a network

model we extend the Rössler variable convention from before to:

$$\begin{aligned} \boldsymbol{x}_i &= (x_i, y_i, z_i)^\top \quad \text{and hence} \\ \mathbb{R}^{\nu \cdot N} \ni \boldsymbol{x} &= \left(\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_N^\top\right)^\top = (x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_N, y_N, z_N)^\top . \end{aligned} \tag{3.33}$$

The basic Equation (2.9) for a networked system was

$$\dot{\boldsymbol{x}}_i = F(\boldsymbol{x}_i) + \sum_{i,j \in \mathcal{N},\ i \neq j} A_{ij} \cdot G(\boldsymbol{x}_i, \boldsymbol{x}_j) \quad \forall i \in \mathcal{N}. \tag{3.34}$$

As the node dynamics should be given by the Rössler oscillator, we identify F with the RHS of Equation (2.14). Furthermore, the linear coupling in the y-coordinate can be formulated by

$$G\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \left(0, q_R \left(y_j - y_i\right), 0\right), \tag{3.35}$$

where q_R is the coupling parameter. Combing these into one set of equations gives

$$\begin{aligned} \dot{x}_i &= -y_i - z_i, \\ \dot{y}_i &= x_i + a y_i + q_R \sum_{j \in \mathcal{N}} A_{ij} \left(y_j - y_i\right), \\ \dot{z}_i &= b + z_i \left(x_i - c\right). \end{aligned} \tag{3.36}$$

In case of this thesis, we use $a = 0.2$, $b = 0.2$ and $c = 5.7$ as parameters, where the intrinsic dynamics of the isolated oscillators is chaotic (cf. Model Description 1). Further, we use $q_R = 0.8$ for which the coupled state is stable. We use two topological structures for the network, that are described in the text when presenting the corresponding example.

The results for each node are presented as the colors of each node in Figure 3.6 and graphed in Figure 3.7 (a) (on a $\log_{10}$ scale). The three generations clearly separate from each other, as can be seen in Figure 3.7 (b), where the lower the generation in the hierarchy, the higher its $T_{RR}^{\mathcal{N}}$ value, i.e. the *later* the node. In contrast, there is no clear correlation to the degree to be found (cf. Figure 3.7 (c)).

3.5.3.2 Random scale-free networks of Rössler oscillators

The next example is an ensemble of 100 random scale-free networks with each $N = 81$ nodes. The networks were created using the Barabási-Albert model of growth and preferential attachment [Bar1999]. The dynamics is again given by the coupled Rössler Oscillators (cf. Model Description 4).

While the number of nodes is the same as for the deterministic scale-free network before, the number of links differs. Before, there were 130 links, while the random scale-free network from the Barabási-Albert model has 158 links. Hence, the results for both examples can be compared qualitatively, but not quantitatively.

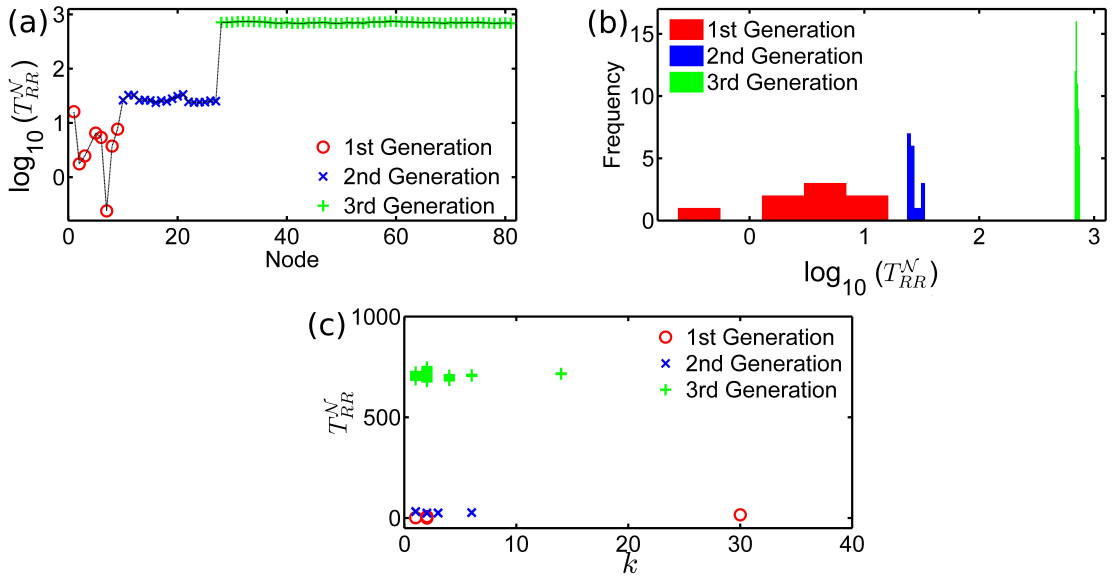


Figure 3.7: $T_{RR}^{\mathcal{N}}$ for for the deterministic scale-free network of coupled Rössler Oscillators. (a) shows the $T_{RR}^{\mathcal{N}}$ resolved per node on a $\log_{10}$-scale. (b) shows the distribution of the $T_{RR}^{\mathcal{N}}$ values and the clear partition in the three generations is visible. (c) The plot compares the $T_{RR}^{\mathcal{N}}$ of each node with its corresponding degree. There is no obvious relation observed. These graphics were published in [Mit2017b].

The distribution of $T_{RR}^{\mathcal{N}}$ for the 100 ensemble of networks structure is shown in Figure 3.8 (a). Further, we computed the expectation value $\langle T_{RR}^{\mathcal{N}}|k\rangle$ of $T_{RR}^{\mathcal{N}}$ given that the node in question has a certain degree k. In Figure 3.8 (b), the striking correlation between $\langle T_{RR}^{\mathcal{N}}|k\rangle$ and the degree k is depicted. In this topology, nodes with high degree have a large influence on the whole network. Hence, an explanation of this correlation might be that the perturbation leads to strong deviations (from the attractor) for the other nodes, too, hence the corresponding trajectory takes a long time to approach the attractor again.

3.5.3.3 Power grid of the United Kingdom

As a more realistic example, we finally investigate a low-complexity model of the United Kingdom's power grid, where the dynamics is given by the coupled Swing equation. The full summary of the model is in Model Description 5.

Model Description 5 (Coupled Swing equation)**.** The coupled Swing equation, also called Kuramoto model with inertia, has been shown to be applicable for grid-resolved modeling of dynamic stability [Aue2016; Hel2016; Men2014].

The state x_i for each node i in the set of nodes $\mathcal{N}$ is described by two variables

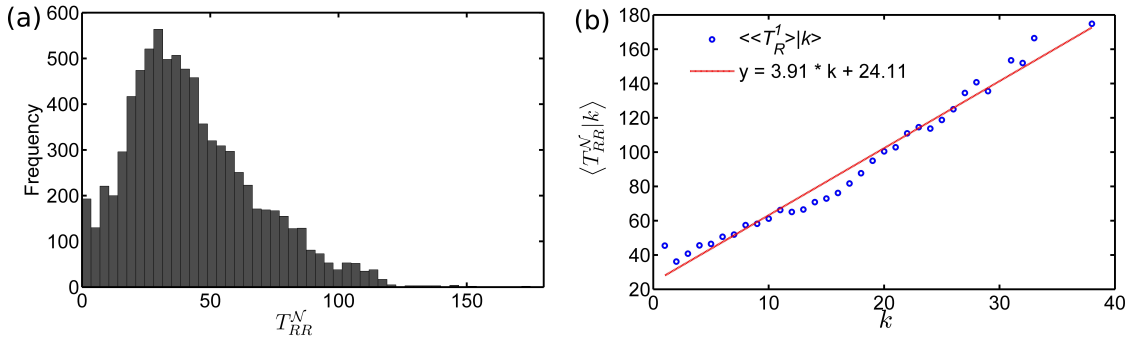


Figure 3.8: (a) shows a histogram of the $T^{\mathcal{N}}_{RR}$ values for all 100 realizations of the random scale-free network of coupled Rössler Oscillators. (b) depicts the striking correlation between $\langle T^{\mathcal{N}}_{RR}|k\rangle$ (the expected $T^{\mathcal{N}}_{RR}$ value condition on the degree of the node) with the degree. These graphics were published in [Mit2017b].

$x_i = (\theta_i, \omega_i)$, where θ_i is the node frequency and ω_i the angular momentum, both relative to the average frequency (cf. the single generator in a power grid in Model Description 3). The oscillators are phase-coupled, meaning their coupling is proportional to the sine of the phase difference. The resulting equations for all $i \in \mathcal{N}$ are

$$\begin{aligned} \dot{\theta}_i &= \omega_i, \\ \dot{\omega}_i &= -\alpha\omega_i + P_i + q_S \sum_{j\in\mathcal{N}} A_{ij} \sin(\theta_j - \theta_i), \end{aligned} \tag{3.37}$$

with the dampening α, the nodal power influx P_i, the adjacency matrix A_{ij} and the coupling strength q_S.

Within the thesis, this model is used to analyze the United Kingdom's power grid. The network consist of $N = 120$ nodes and 165 transmission lines, i.e. links, that are encoded in the adjacency matrix A. Furthermore, we randomly chose $\frac{N}{2} = 60$ nodes to be a consumer with $P_i = -P_{ref}$ and the other half to a producer with $P_i = +P_{ref}$. The parameters have been chosen to be $P_{ref} = 1$, $\alpha = 0.1$ and $q_S = 9$. The topological structure of the network is depicted in Figure 3.9.

The results of the computation of $T^{\mathcal{N}}_{RR}$ for the UK power grid are in Figure 3.9, where the nodes are colored corresponding to their $T^{\mathcal{N}}_{RR}$-value. Summarized in Figures 3.10 (a) and 3.10 (b), there are some clear outliers (marked in red). A relation to the degree k or betweenness bc (see Section 2.1 for definition) has been analyzed but could not be found (cf. Figures 3.10 (c) and 3.10 (d)).

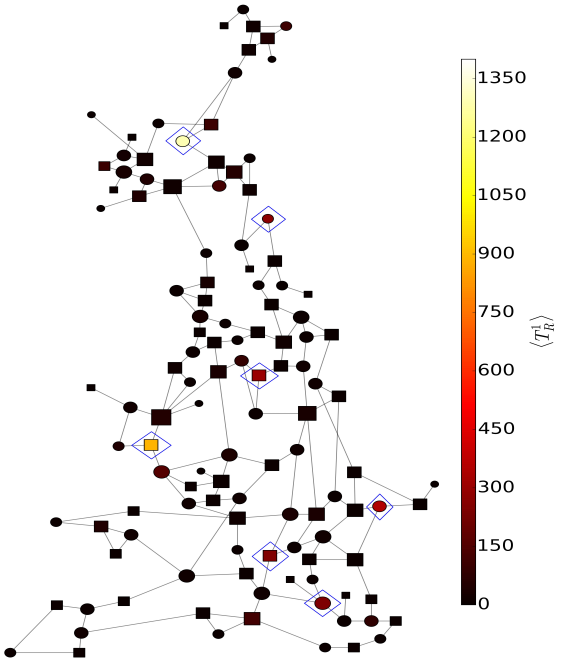


Figure 3.9: The United Kingdom's power grid is composed of $N = 120$ nodes. Their coloring is the result of the Timing of Networks analysis using the coupled Swing equation as dynamics (cf. Model Description 5). There are 60 generators (circular nodes) and consumers (square nodes) and they are distributed randomly. The size of each node is proportional to its degree. The nodes encircled by blue diamonds are particularly *late* nodes. After a perturbation, they may go through long transient phases before returning to the attractor. Hence, there is damage to the machines possible and particular care of the operators is advised. This graphic was published in [Mit2017b].

3.6 Discussion

In order to see how far the two proposed metrics answer the question "How can we properly quantify the time to reach a system's attractor?" we will go along the four essential conditions that have been worked out in Section 3.2 for this discussion: (I) finite reaching time, (II) physical interpretation (III) continuity and (IV) invariance.

Area under distance curve. D has been defined as the cumulative distance to the attractor over time in order to emphasize the idea that a trajectory stays "far" from the attractor in the transient while being close in the asymptotics. In that sense, it can be understood as *time weighted by the distance* (to the attractor). The distance $d(\cdot, \cdot)$ is not necessarily meant in the mathematical sense [Hei2002], but it only needs to approach 0 around the attractor and be 0 on it. In that way, it is possible to choose the appropriate $d(\cdot, \cdot)$ for different research questions, e.g. asking about costs or damages. Even in these interpretations D is a metric capturing the transient time, because there are only

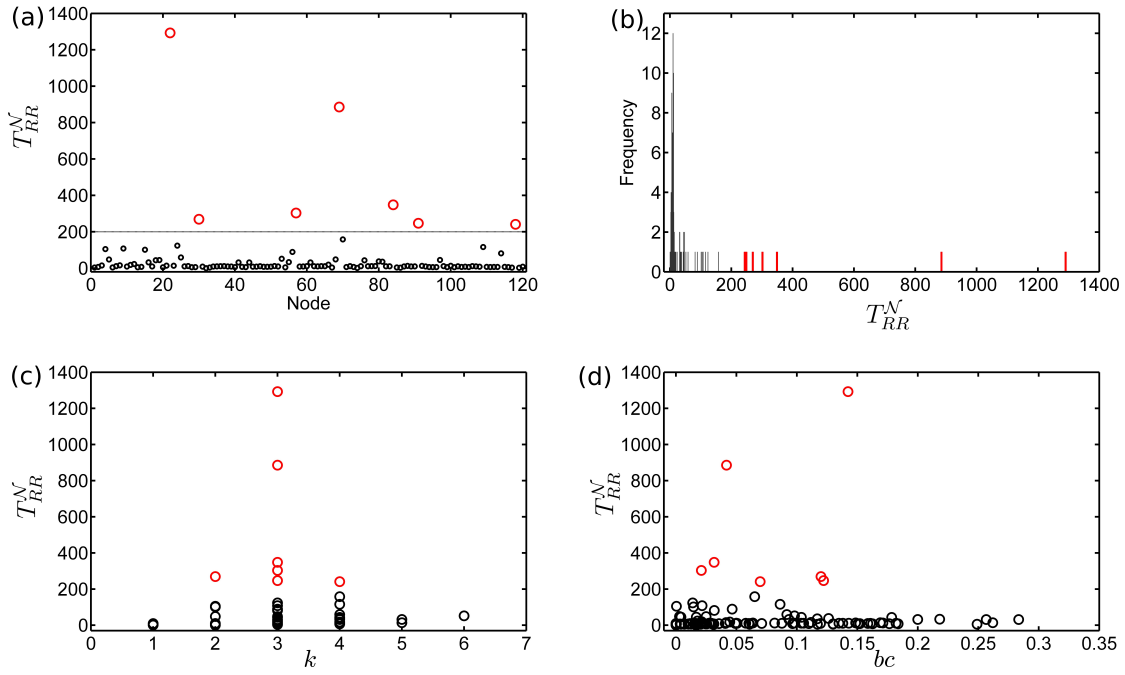


Figure 3.10: (a) shows the $T_{RR}^{\mathcal{N}}$ value for all the $N = 120$ nodes of the United Kingdom's power grid using the coupled Swing equation (cf. Model Description 5) as dynamics. (b) is a histogram of these results, showing how 7 nodes have a distinctly higher $T_{RR}^{\mathcal{N}}$ value than the others ($T_{RR}^{\mathcal{N}} > 200$, marked in red). (c, d) portray the relation of $T_{RR}^{\mathcal{N}}$ to the degree (k) and the betweenness-centrality (bc) respectively. for each node with their Dependence of $\langle T_R^1 \rangle$ on (c) degree (k) and (d) betweenness centrality (bc) of the nodes. The nodes with distinctly higher values of $T_{RR}^{\mathcal{N}}$ are the very *late* nodes. Hence, a perturbation at this node may lead to a very long transient before returning to the attractor and could potentially lead to damages of the machines. These graphics were published in [Mit2017b].

contributions when the trajectory is distant from the attractor, i.e. still in its transient phase. Another point of view is to see D as the time to reach the attractor weighted by the distance.

We understand Condition I, finite reaching time, as fulfilled. For hyperbolic attractors and $d(\cdot, \cdot)$ being a mathematical distance function, the integral in Equation (3.5) does converge. Trajectories approach the attractor exponentially in the asymptotics and the integral over the exponential envelope is finite.

While this covers most systems relevant for real-world applications, in some very specific cases, D might be infinite. The asymptotic tail of the integral might not converge, i.e. the trajectory does not approach it "fast enough". This means either this is the desired result or $d(\cdot, \cdot)$ has not been chosen appropriately. In the first case, it could be for example that D was computed for an initial condition that is not economically feasible, so the cost diverges. This would imply that even though the attractor is systemically

stable, it is not economically feasible to cope with small perturbations.

From a technical perspective a divergence in D can be understood as indicating that d has not been chosen matching to the system. For instance, using the Euclidean distance and $\dot{x} = -\frac{1}{2}x^3$ where the solutions are $\varphi(t, x) = \text{sign}(x)\frac{1}{\sqrt{t+|x|^{-2}}}$, D does not converge. Another example is to take a linear system $\dot{x} = -x$ with $x < 1$. Using $d(x, \{0\}) = -\frac{1}{\ln|x|}$ with $d(0, \{0\}) = 0$ gives $D \to \infty$.

This can usually be solved by choosing an appropriate d. For example, choosing for $\dot{x} = -\frac{1}{2}x^3$ using $d(x, \{0\}) = \exp(-|x|^{-1})$ and $d(0, \{0\}) = 0$ gives finite values for D.

Condition II is fulfilled because there is no direct parameter. Still, as there is the indirect dependence on d a discussion is necessary and given in comparison to the first and last entry time to an ϵ-environment $T_F^\epsilon(x)$ and $T_L^\epsilon(x)$ respectively. For them, a small change in ϵ will have a huge impact on the measured times because for $\epsilon \to 0$ both values go to infinity. Furthermore, if one would locally change the way how the distance to the attractor is measured, the values for $T_F^\epsilon(x)$ and $T_L^\epsilon(x)$ would change drastically, too. On the contrary, as D is defined as the cumulative d over time, a local change in d will have only minor effects on the exact value, so even estimated functions for d with some uncertainty can be used.

Condition III, continuity, has been implemented in D by using the integral representation. Hence, the function is even differentiable along the flow (see Equation (3.6)).

We see Condition IV, invariance, as fulfilled, if d has been chosen with some meaning, e.g. economic damages. Then one can simply represent the economic damage function in the changed coordinates, because the meaning is independent of the coordinates. This reasoning is not mathematical but context-dependent. From a purely mathematical point of view, if d is just any distance function, generally the result is not invariant under change of coordinates as it depends on geometric features of the system. But as we want to go in the direction of real-world systems, a model-specific choice of d is compulsory anyway.

Regularized reaching time. T_{RR} has been defined as the difference in time to approach the attractor with respect to a reference trajectory starting at x^{ref}.

Condition I, finite reaching time, is fulfilled everywhere for all states within the basin of attraction except the attractor and the strong stable manifold. In case of the attractor, the trajectory will stay on it while trajectories from other points approach the attractor only; and, by definition, points on the strong stable manifold approach the attractor a lot faster, also in the asymptotics. So these infinities are actually reasonable results. Also, both are usually of a smaller dimension than the state space. Hence, the probability of choosing such an initial condition is 0 (assuming distributions that are not Dirac delta function-like), and these cases are rather irrelevant for real-world applications.

Conditions II and III are intrinsically fulfilled by avoiding parameters. The necessary choice of x^{ref} introduces a constant shift only while not changing the structure of the function. When looking at central moments of T_{RR}, i.e. ones invariant under shifts, this dependence on the choice of x^{ref} disappears completely as they would only shift the mean. So an analysis by changing the system's parameters is possible. This has been done in

the examples for the global carbon cycle and Rössler system and T_{RR} has been confirmed as an early-warning signal. This analysis can be seen as a systemic approach to the concept of *critical slowing down* (CSD) [Len2011; Sch2009; Sch2012] after a shock, i.e. an instantaneous and non-infinitesimal perturbation, uncovering prebifurcational changes in the transient behavior. In contrast, CSD is usually done with (local) noise only. The usage of shocks has been developed in the context of Basin Stability [Men2013; Men2014] and its extensions [Hel2016; Kan2016; Kli2015; Mit2015; Mit2017a].

Condition IV, invariance, is proven to be fulfilled for hyperbolic fixed points. In case of more complex attractors, we can currently only define an estimation of T_{RR} which depends on geometric properties. So invariance might not be given and more research is due in that direction. An important step in that direction has been done by writing down the properties of Δ which imply that the necessary way of measuring how a trajectory approaches might not be local (except fixed-points). The used pragmatic choice of $\Delta = D$ demonstrates this as it basically says that the remainder of the trajectory should have an ϵ small value of D only.

An assumption that has been made during the proof of invariance of T_{RR} for hyperbolic fixed points is: the eigenvalue of the RHS's Jacobian with the largest real-part is either unique and with multiplicity 1 or there are two that are complex conjugated to each other. However, this condition is not really constraining because we assume most real world systems fulfill it.

Comparison. The metrics have been applied to several examples and we will discuss a comparison between both metrics here. They are depicted in Figures 3.3 (b), 3.3 (e), 3.3 (h) and 3.3 (k) and show different relations, as stated in the figures. The exponential lower bound and the exponential relation for the linear system and swing equation respectively come mostly from the asymptotic behavior, in particular as the linear system does (by definition) not have any nonlinearities. Still the relations are different as the asymptotic behavior differs slightly, too, one being a node the other a focus. This shows that even though we clearly focus on the transient, it is actually important to be aware of the asymptotic behavior, too. And one cannot analyze the former without knowing about the latter.

In contrast, the linear relation for the global carbon cycle really points to the transient behavior only. It is due to the states passing by the "desert-like" saddle. Finally, there seems to be no clear relation between both metrics for the Rössler example, pointing to the chaotic behavior. Still, both metrics have separately been useful, D demonstrating the smoothness of the basin of attraction and the standard deviation of T_{RR} being sensitive to qualitative changes of the system.

Regularized reaching time as a network metric. $T_{RR}^{\mathcal{N}}$ was based on the definition of T_{RR} and extended it to a network measure. It associates to each node how much later (on average) the system returns to the original attractor after a perturbation compared to the earliest node. For the coupled Rössler oscillators, we have seen a clear separation of the $T_{RR}^{\mathcal{N}}$ values into the generations of the deterministic scale-free network, but no relation

with the degree. In contrast, for the random scale-free network, we could observe a striking correlation between the $T_{RR}^{\mathcal{N}}$ and the degree This result can probably be explained with high-degree nodes having a strong influence on the others, particular in a scale-free network. But this does not explain the lack of this correlation in the deterministic scale-free network. This results implies a (future) question: What are the (network-)topological difference between these models leading to this difference? The third example was to use the coupled Swing equation on the United Kingdom's power grid. Some nodes with very high $T_{RR}^{\mathcal{N}}$ values have been identified. After perturbations, these will go through long transients and might lead to damages of the system along the way. Hence, an investigation should be done on how to decrease their $T_{RR}^{\mathcal{N}}$ value.

Other Methods. When developing this research on measuring times to approach the attractor, we had the impression that there are two more common ideas, additionally to the first and last entry time. We do not intend to have a complete overview of all methods but would like to discuss these two shortly here. This part refers to a general system in the sense of Equation (2.6).

The first idea is to develop metrics based on characteristic times. These are usually defined as the time until a quantity is reduced to $1/e$ of its original value [Cla2011]. This quantity could be a distance to the attractor or a coordinate. From this definition it already follows that they are subject to problems concerning Condition IV. Also, even if the quantity is at $1/e$ of its initial value, the trajectory might still be far away from the attractor and in its transient dynamics. Lastly, taking a one-dimensional linear system and assuming the quantity is the coordinate, the characteristic time is constant for all initial conditions. This is counter-intuitive when thinking about a time to approach the attractor.

The second idea for general systems is to use Lyapunov exponents [Cvi2016]. They have units of inverse time and are invariant under changes of coordinates. However, they are actually a property of the attractor. So they do not capture the transient away from the attractor, but only the asymptotics closely around it.

3.7 Summary

In the context of transient analysis of complex dynamical systems, we have treated the question in this chapter: "How can we properly quantify the time to reach a system's attractor?"

First, we have worked out the four essential conditions of quantifying the timing of transients in order to develop two new metrics, *Area under Distance Curve D* and *Regularized Reaching Time* T_{RR}. As the focus of this work is meant to be on making a first step to real-world systems, we have applied the metrics numerically to 4 chosen examples systems, observing different features. Finally, we have discussed in detail how far the metrics treat the four essential problems.

With this approach, new features of the examples have been uncovered. In the global

carbon cycle model, we have demonstrated the importance of the transient analysis, as the desert state is only a saddle but nevertheless passing by there would lead to an extinction of humanity. The splitting of the basin of attraction is partially due to the strong stable manifold of the attractor. But it continues for lower values of $c_{terrestrial}$ where it is only due to quantitatively different behavior demonstrating the need for quantitative methods. Particularly interesting is how the (central) statistics of our metrics are a systemic approach to the concept of *critical slowing down* leading to an interpretation as early-warning signals, which we have demonstrated also. The independence of the choice of reference points has been achieved by the usage of central moments.

In case of the generator in a power grid, most of the relevant dynamics seems to be dominated by the linearization of the equations around the focus.

In order to prove the applicability to more complex dynamics, we have used our metrics on the Rössler system, too, and found the smoothness of the attractor's basin with D. As the attractor itself is chaotic, this smoothness is surprising. T_{RR} reacts strongly to the sensitivity to initial conditions of the chaotic system. Still, its worth is displayed when varying the a parameter. This parameter has strong influence on the Rössler system's dynamics and T_{RR} reacts strongly to the different bifurcations and even the chaos-chaos transitions, proving again its worth as *early-warning-signal*.

The detailed discussion on the two metrics have showed that, while they do treat the four essential problems, they do not fully solve them and further investigation is needed. Also, they come from two very different basic ideas so the comparison showed that they really measure independent features but can improve the understanding of a system by combining them. For both metrics, we have showed that they are Lyapunov-functions. While some properties have already been used in this thesis, these definitions in terms of orbital derivatives may be a rich groundwork for the next steps.

We have not performed any comparative analysis with the mentioned first- and last-entry-time approaches because these behave inconsistently and their quantitative results are arbitrary, as discussed at length in Section 3.2.1.

Chapter 4

Topology of Sustainable Management

The second main question mentioned in Chapter 1 *Introduction* "Can we avoid transgressing boundaries and stay safe (& just)?" is deeply rooted in the research for a sustainable future of humanity. The idea is that if humanity does not respect certain boundaries, a prosperous future for all of us might not be possible. Within this context, different notions of boundaries have been proposed recently. Prominent among them is the idea of *Planetary Boundaries* (PBs) developed in [Roc2009] and extended in [Ste2015b]. They define a set of biophysical boundaries chosen in order to enable a sustainable human development if they are respected. While they provide safety from biophysical degradation of the Earth system by defining a *safe operating space*, the social component of humans is not explicitly included. Raworth [Raw2012] pointed out that a human future needs to concern the social component as well, demanding for *Social Foundations* (SFs). The collection of all situations where humanity respects both of these concepts, the PBs and SFs, is called the *safe and just operating space*, often referred to as SAJOS. Other discussions on such kinds of constraints are within the frameworks of the United Nations Sustainable Development Goals [Div2017] , the "tolerable environment & development (E&D) window" and the "guardrails" in a model of the Earth system [Sch1998].

Currently there is much research on refining the current definitions of the individual PBs, e.g. for freshwater [Ger2013] and phosphorus [Car2011], extending them, e.g. terrestrial net primary plant production [Run2012], and downscaling them from the global to a regional level [Häy2016]. But another direction of research has emerged recently: their "interaction" due to the system's intrinsic dynamics [And2013; Hec2016].

In this context, a mathematical framework to analyze complex systems with boundaries and the inclusion of the dynamics called *Topology of Sustainable Management* (TSM) was developed [Hei2016]. On a set of states, i.e. a state space, we assume a *default dynamics* and options for influencing the dynamics that we call *management*. Furthermore, we consider a distinction of the state space into an undesirable and a desirable region. These are seen as external input for the analysis and could be given by the PBs or SFs. The result is a partition of the system's state space in different regions, complex structures in state space corresponding to a hierarchy of safety levels. These regions are a *qualitative* classification and differ in how secure they are and how much management one needs to

either stay within the desirable region or reach it.

The importance of such a qualitative analysis before a quantitative optimization of a pathway to a sustainable future becomes obvious with certain *dilemmas* naturally arising in this framework. Here, a dilemma is a situation where a qualitative choice is present. In particular the so-called *lake dilemma* will play an important role and is elaborated in this chapter in detail.

In [Hei2016], a dynamical system was defined as a family of trajectories on a state space with some basic topology (hence the name of the framework). Further, we have carefully chosen the term *management* in order to emphasize that usually only a slight influence on the dynamics might be possible but not a complete control, e.g. in the global climate system. Still, this is formally equivalent to a control system as shown in [Kit2017e] and I will follow this control-based point of view in this thesis. Thus, we assume the system to be based on a deterministic ordinary differential equation with a control parameter, i.e. a *control system*, leading to what we call the *variant definition of TSM*. This enables us to draw the connection to Viability Theory (VT) and makes the tools from that field available. In particular, we will make use of the so-called Saint-Pierre algorithm (SPA) to enable the step from manually analyzed two-dimensional examples to an automatically analyzed three-dimensional model. We call the latter step the *operationalization of TSM*.

VT was developed by Aubin and his collaborators [Aub1990; Aub1991; Aub2011; Aub2012] as a subfield of control theory, addressing the problem of maintaining a system's trajectory within a *constraint set*, e.g. delimited by planetary boundaries. The main concept, the *viability kernel*, is defined as the set of states from where it is possible to keep its trajectory within the desirable set indefinitely. Hence, it is possible to avoid the transgression of the boundaries. The second main concept is the *capture basin*. It is defined with respect to a *target set*, i.e. a chosen set of states that one wants to reach. Then, the capture basin is the set of states from which the target set can be reached while not leaving the aforementioned constraint set. Hence this concept is very close to "reachability" in control theory. VT has been applied in many domains, e.g. economics [Aub1997], fisheries [Cha2008; Cur2005], energy consumption of wireless sensors [Kon2017], forest management [Mat2015], language competition [Cha2010], social networks [Mat2017b], sustainability management [Del2008; Rou2013; Rou2014], resilience modeling [Def2011; Mar2004]. In particular, VT was recently used for showing our rapidly shrinking capacity to comply with the planetary boundaries on climate change [Mat2017a].

Within VT, several algorithms for the estimation of viability kernels have been developed, among them the Saint-Pierre algorithm (SPA) [Sai1994] with extension to machine-learning-based classification methods [Def2007], reachability based viability [Mai2013] or optimal-control-problem-based ideas [Bok2006; Del2008; Rou2013].

We extend this work by developing a *nonlinear, local time homogenization* solving the problem of vastly differing time scales when estimating viability kernels and capture basins using the SPA. It fully homogenizes the time scale of dynamics while keeping the major properties of the system invariant. Even though the approach is straightforward, it is essential for the analysis of the three-dimensional example model.

Another extension deals with the problem that commonly the state space of a model can be unbounded in at least one variable, e.g. economic production. But the SPA is for bounded sets only. To tackle this problem of unbounded state spaces, we propose a nonlinear coordinate transformation to a bounded state space. In particular, this transformation has the property that one can choose a scale to be "resolved best".

Certain regions of the TSM-partition, the so-called *eddies*, are shown to be naturally defined as implicit target sets. To identify these automatically, we also have developed an extension of the SPA.

With the operationalization of TSM working, the door is open to a systematical, numerical bifurcation analysis. With that we mean qualitative changes in the structure of the TSM-partition when varying a parameter of the model. We will give a first idea of that for a three-dimensional example model. Inspired by the work on Basin Stability in [Men2013; Men2014], where the volume of the basin of attraction has been taken as an indicator for stability, the volume of the different regions within the TSM-partition has been found to be a valuable indicator for the bifurcation analysis.

The chapter is structured as follows. After this introduction follows a short section on software availability. Then we continue with a metaphorical description of the TSM-framework without equations in Section 4.2. As this work is meant to be in the interdisciplinary context of sustainability, we find it important that the basic concepts are understandable without specific mathematical knowledge. Using the basic definitions of VT introduced in Section 2.3, we present the variant definition of TSM in Section 4.3. This framework definition is illustrated with multiple 2-dimensional examples in Section 4.4. Then, we give a description of the Saint-Pierre algorithm and introduce the extension that were mentioned before in Section 4.5. Combining the work from before, we analyze a three-dimensional model, that we call the AYS-model, in Section 4.6, including a bifurcation analysis in Section 4.6.8. Finally, we summarize and discuss the results in Section 4.7

This chapter is based on the publications "Topology of sustainable management of dynamical systems with desirable states: from defining planetary boundaries to safe operating spaces in the Earth system" [Hei2016] and "Operationalization of Topology of Sustainable Management to Estimate Qualitatively Different Regions in State Space" [Kit2017e], where I co-authored the first work and lead-authored the second.

4.1 Software Availability

For this project, an open source Python-library for viability computations called `pyviability` has been developed and is available online [Kit2017c]*. The source code for the model-specific computations of the AYS-model (cf. Section 4.6) including a description on how to reproduce the results is available as well [Kit2017b]†. In case of questions, please

*`https://timkittel.github.io/PyViability/`

†`https://github.com/timkittel/ays-model/`

contact the author of the thesis.*

4.2 Verbal Description of TSM

Before going in all mathematical details, we found it important to give a verbal description of TSM. This is meant to provide to the idea of the TSM-framework to a reader that is not inclined to go through all the mathematical details of this work.

For the explanation, we take the common metaphor that "we're all in the same boat" literally and represent the state of the Earth system with all its natural and socio-economic parts by a boat on a rather complex system of waters such as in Figure 4.1.

The boat can only be on water, not on land, will generally float along with the stream that represents the inherent dynamics of the Earth system over hundreds and thousands of years (the "default trajectory"). The person on the boat, representing humanity, may also row against the stream if he is not within the rapids (marked by swirls). This possibility of rowing represents mankind's agency in deliberately influencing the Earth system's. "Floating along the river" will be called *default dynamics* and "rowing against the stream" will be *management*. Further, we assume a qualitative distinction with regard to where humanity wants their boat to be. This is represented by the division of the whole picture in a sunny, *desirable region* (left) and a dark, *undesirable region*† (right). The desirable region is meant to consist of all those possible states of the natural and socio-economic parts of the Earth system in which some generally agreed environmental and living standards are met, e.g. the aforementioned PBs [Roc2009; Ste2015b].

While the state space of the Earth system may be highly complex, one may generally assume that all of the waters explained in the follow exist within the eath system's state space as well.

There will be a certain sunny water region where one does not need to row at all to stay in the sun forever but can simply lean back and let the boat float around inside that region. In Figure 4.1, this region is the top-left tranquil tarn, that we call *shelter* to emphasize their desirable and safe nature. Indeed, we will argue below that these shelters may be the most natural candidates for being called a "safe and just operating space for humanity" (SAJOS), only that we may not yet be inside. In the Earth system, there may be several such shelters, one of which might correspond to resilient states of the world [Fol2010] where humanity lives reconnected to the biosphere [Fol2011] and no active intervention or constant large-scale management is needed.

Connected to the shelters, there will be parts of the desirable region where there is a need for rowing, i.e. management, to reach the shelter. Else one might be drawn into the dark. A region where one can reach the shelter without entering the dark in-between is

*Tim.Kittel@pik-potsdam.de

†Interestingly, my stay in Brazil as an exchange within the IRTG 1740 "Dynamical Phenomena in Complex Networks" taught me a strong cultural bias we had when developing the whole nomenclature. While in Germany, "sunny" is often associated with good and "dark" with bad, in Brazil "sunny" is associated with the scorching and burning sun and "dark" with the cool and relaxing shade.

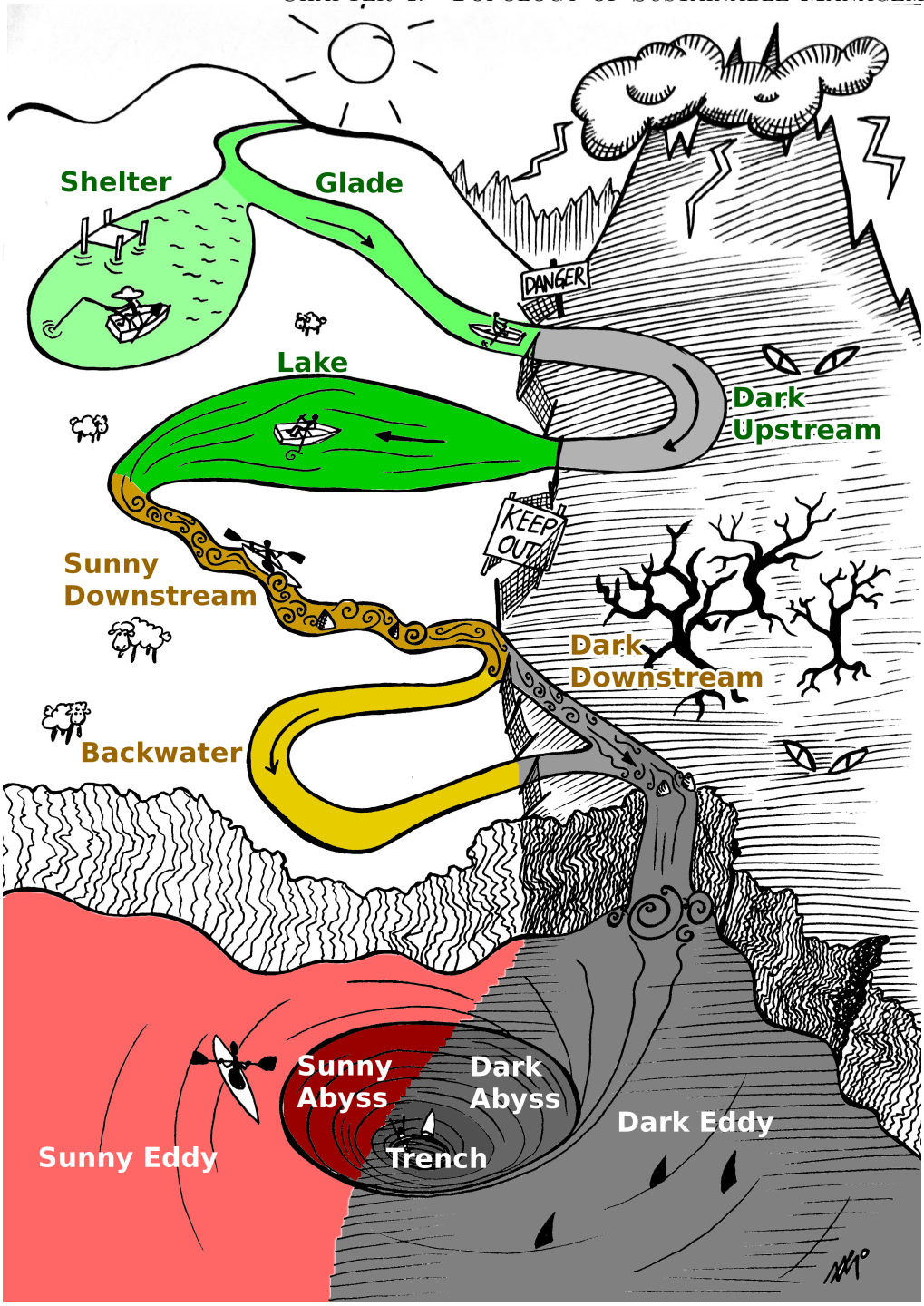


Figure 4.1: A metaphorical model of boat on a river flowing from the mountains to the sea used for the verbal description in Section 4.2. The default flow is downwards the river, except in the shelter where there is no default movement. A person in a boat on the river can decide to flow along the stream (default dynamics) or row against the stream (management), except within the shelter where the default dynamics is to stay calmly there. Within the fast parts (rapids) marked by swirls, no rowing against the stream is possible. Each point on water corresponds to a state, hence the whole waters are the state space. The left part of the picture is the desirable region while the right part is undesirable. This model has been chosen to give rise to all regions except the dark upstream, such as *glades* and *lakes* from where the shelter can be reached, *backwaters* from where one can at best stay in the safe by management, and several types of worse regions, all labeled here and explained in the text. See also the decision tree in Figure 4.2 and a different metaphorical hand-drawn picture in Figure 4.3, that is focussing on the up- and downstream. This graphic was published in [Hei2016].

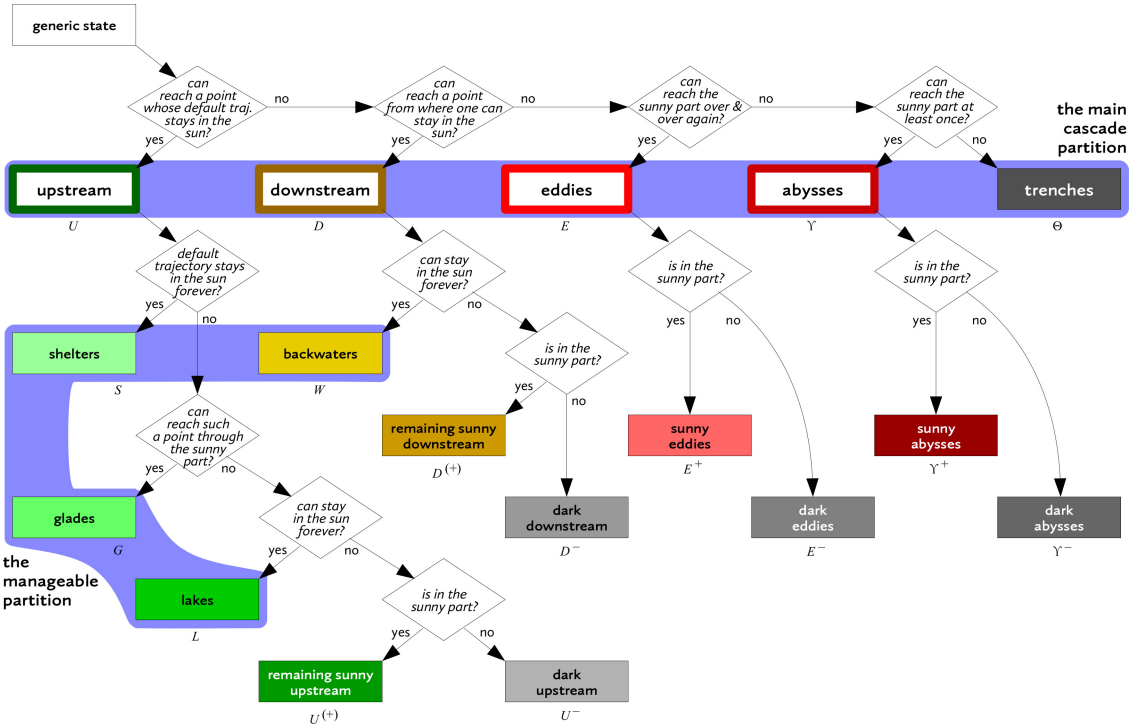


Figure 4.2: Decision tree summarizing how for a single state one can find out to which region of the partition it belongs. The first row of question introduces the distinction within the main cascade into upstream, downstream, eddies, abysses and the trench. Within each, a finer distinction is done by additional questions, except for the abyss. Further, the manageable region is marked, as it overlaps with the up- and downstream. The color scheme has been carefully chosen and is used for the remaining pictures to ensure understandability and consistency. This graphic was published in [Hei2016].

called a *glade* to emphasize their "almost safe" character. In Figure 4.1, the corresponding region is the part of the river passing by the shelter.

Another part of the sunny region is called a *lake*. From here, the shelter is reachable by passing through the dark only, but one may stay inside the sunny region forever if one rows continuously. Inside the picture, this corresponds to the water above the "keep out" sign. This gives rise to the so-called *lake dilemma* of whether to stay in the sun for ever but being in continuous need of management or risking a temporary passage through the dark but finally reaching the shelter. Below is a more detailed explanation. All other regions from where one can reach the shelter are called the *remaining sunny upstream* if in the sun, or the *dark upstream* if in the dark. The combination of all these regions is the *upstream*, as from each point within the shelter is reachable.

Further down the stream there are places where it is still possible to stay in the sun forever, only that one has to row over and over again to do so. In Figure 4.1, this corresponds to the slow-moving side branch below the "keep out" sign. Such regions, called *backwaters*, are similar to lakes, only without the option of rowing to a shelter.

Hence, the lake dilemma does not occur. While the upstream was defined by being able to reach a shelter, the *downstream* is now defined as all places from where a backwater but not a shelter can be reached, including the backwaters. An example for a backwater could be a "machine world" where humanity can fully control nature to its very minute detail. While they can stay within the sunny region for infinite time by this management, there is no way of reaching a shelter anymore because the ecosystem has been changed irreversibly.

While in the upstream and downstream, one may finally stay in the sun forever, less hopeful places exist (below the waterfall in Figure 4.1), too. In the *eddies*, one can make sure to always reach the sun, but will be forced into the dark over and over again. The *abysses* are similar, but one may reach the desirable part only a number of times and then finally end up in the *trench*. Within the trench, there is no possibility to escape the dark.

This completes our main partitioning of the Earth system's or any other manageable system's state space into qualitatively different regions: Upstream and downstream defined by being able to reach shelters or backwaters, abysses defined by not being able to avoid ending up in a trench, and eddies in between, defined by being at least able to switch between sun and dark forever. Figure 4.2 summarizes all these regions in the form of a decision tree, where one can identify the region the system is in by answering a small number of questions.

The lake dilemma was mentioned before to arise within a lake. It is the most notable among the five dilemmas identified in [Hei2016] and the one relevant for this thesis. Thus, we focus only on this one and refer the reader to the paper for more details.

We call a *dilemma* a situation where a qualitative decision needs to be made. In other words, instead of optimizing how to reach a certain goal, one has two qualitatively different goals at hand that may not be comparable.

The *lake dilemma* is the choice between (a) uninterrupted desirability (that is, not transgressing the boundaries) or (b) transgressing the boundaries temporarily but finally reaching a region where one does not have to apply management anymore.

An example, where a lake dilemma can occur, may be the following: A transformation of the global energy system towards a carbon-free economy might temporarily lead to welfare losses in poorer countries. So one may choose not to follow through with this transformation but instead stay with a partially fossil-based economy, but is in constant need to ensure the effects of the carbon emissions are taken care of. Still, the temporary welfare loss in poorer countries, e.g. due to famines, is avoided. Both a viable choice and qualitatively different.

A particular case of the lake dilemma is the *pressing lake dilemma*, where the above decision has to be made within a finite, i.e. limited, time. That means this situation is aggravated as the decision cannot be pushed back but has to be made until a fixed deadline.

Picking up the example from before, this could happen if due to the partially fossil-

based economy a lock-in effect arises. Then, there might be a point in time after which one may not be able to change the path to a carbon-free economy anymore.

4.3 Variant Definition in Terms of Viability Theory

As motivated in the introduction of this chapter, we will give a variant definition of TSM based on the ideas of Viability Theory presented in Section 2.3.

We elucidate some central notions using a metaphorical example of ducklings in Figure 4.3. The water region represents the state space and the streamlines represent the dynamics. The ducklings can either swim with the flow (default dynamics) or struggle and swim against it (management). However, it is not possible to swim up a waterfall once they have dropped down. The desirable region, the safe environment providing enough food and nesting places for the ducklings, is on the left and the undesirable region full of predators on the right. In the following, we will always introduce the general concept for each region first and then explain where it comes up in the ducklings example*.

The definition is based on a general control system (Equation (2.15), here only repeated)

$$\dot{x} = h\left(x, u\right), \tag{4.1}$$

where we additionally require that a *default control* $u^{\mathrm{d}} \in \mathcal{U}$ is separated out from *all possible controls* $\mathcal{U}$. $\mathcal{U}^{\mathrm{m}} = \mathcal{U} - \{u^{\mathrm{d}}\}$ are the *management controls*.† Hence we call $h(x, u^{\mathrm{d}})$ the *default flow* giving rise to the *default dynamics* and the dynamics corresponding to the manageable controls is called *management options*. Furthermore, we require a division of the state space into *desirable* $\mathcal{X}^+ \subseteq \mathcal{X}$ and *undesirable* $\mathcal{X}^- := \mathcal{X} - \mathcal{X}^+$.

When approaching the idea of TSM from VT's point of view, we can found it on two basic quantities, the *shelters* $\mathcal{S}$ and the *manageable region* $\mathcal{M}$

$$\mathcal{S} := \mathrm{Viab}_{\{u^{\mathrm{d}}\}}\left(\mathcal{X}^+\right) \tag{4.2}$$

$$\mathcal{M} := \mathrm{Viab}_{\mathcal{U}}\left(\mathcal{X}^+\right), \tag{4.3}$$

where the definitions of the viability kernels can be found in Section 2.3. Shelters are the viability kernel of the default dynamics, meaning the system will stay in the desirable region forever *without management*, thus being the safest regions in the state space. For the ducklings the shelter is the nest, a place in which they can stay in a safe environment forever without swimming against the stream. Inside the manageable region, one can also

*The change of the picture from Figure 4.1 to Figure 4.3 is not 100% necessary here. Still, there are some minor differences that were the reason to develop a second picture for the second paper. First, Figure 4.1 does not contain the distinction between the time-limited und time-unlimited lake. Second, Figure 4.3 focuses more on the upstream regions which are the most relevant for the analysis of the AYS-model in Section 4.6. Finally, I found both pictures to be very aide-memoire and hope this is similar for the reader.

†Here and in the following we use the lax difference and disjoint union notation with "−" and "+" for sets.

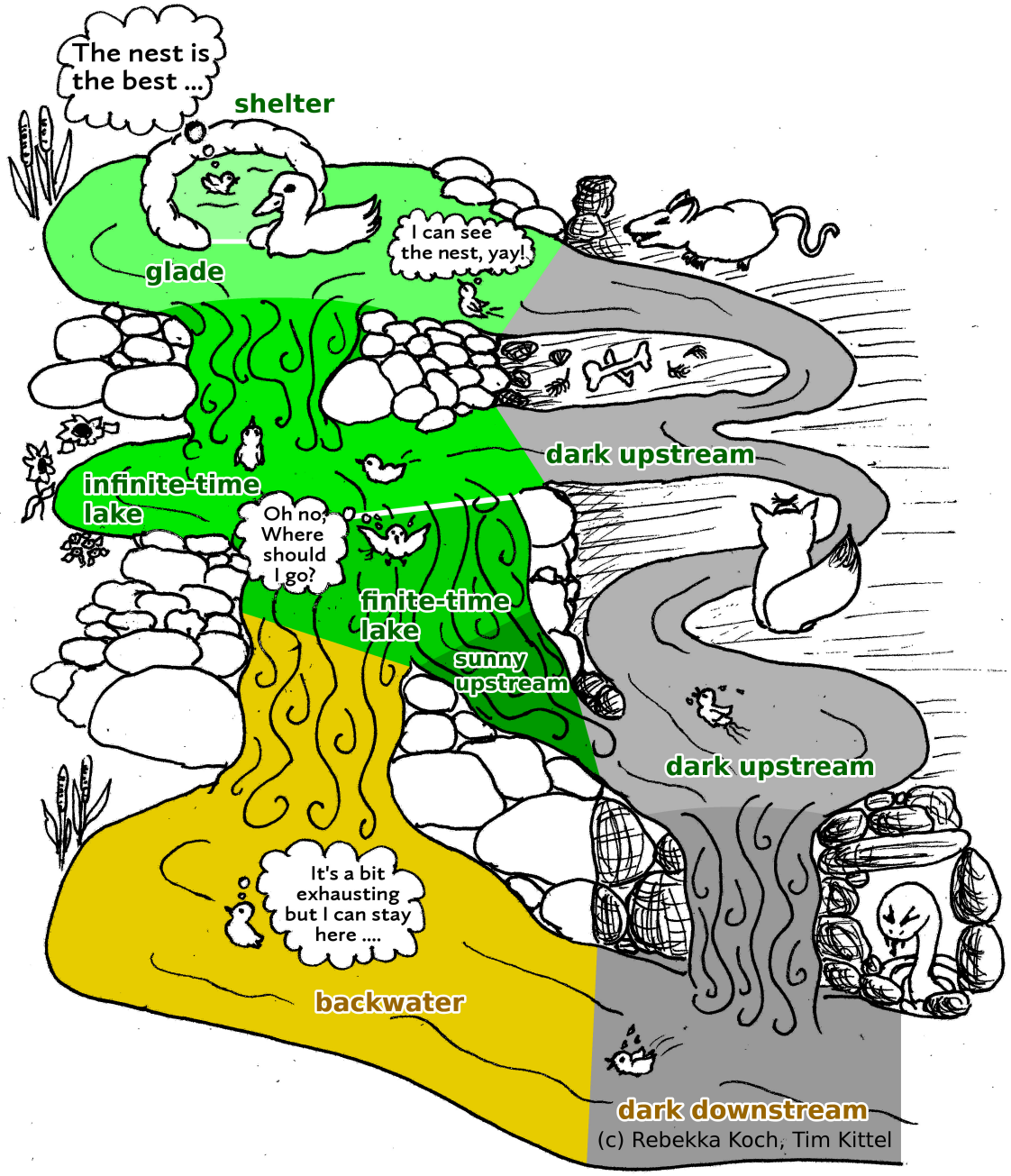


Figure 4.3: Metaphorical example for the concepts introduced in Section 4.3, which is a variant definition of the Topology of Sustainable Management (TSM) framework developed in [Hei2016]. It depicts a river flowing downwards with ducklings on it that may swim through the desirable (left) and undesirable parts (right). In the nest, which corresponds to the *shelter*, the ducklings do not have to swim and can stay there forever without effort. Outside they will slowly drift down or can swim, i.e. *manage*, against the downward-flowing stream, if the the stream lines are not curly, but long. In areas with curly stream lines, the waterfalls, the flow is so strong that the ducklings move with the flow anyway. The example gives rise to a number of qualitatively different regions from the TSM-framework. This graphic was published in [Kit2017e].

stay in the desirable region forever, but may need *all* possible dynamics. In the case of the ducklings example it corresponds to the waters in the desirable region in Figure 4.3 (including the shelter). This is a special case only as we did not want to make the picture too complex and we refer the reader to the detailed explanations in [Hei2016] for the full framework.

The set $\mathrm{Capt}_{\mathcal{U}}(\mathcal{M})$ from where the manageable region can be reached is naturally divided into the *upstream*

$$\mathcal{U} := \mathrm{Capt}_{\mathcal{U}}(\mathcal{S}) \tag{4.4}$$

from where the shelter is reachable and the rest, called *downstream*

$$\mathcal{D} := \mathrm{Capt}_{\mathcal{U}}(\mathcal{M}) - \mathcal{U}. \tag{4.5}$$

As both sets, the upstream and the downstream, may have qualitatively very different dynamics within, we introduce a finer partition. From the *glades*

$$\mathcal{G} = \mathrm{Capt}_{\mathcal{U}}^{\mathfrak{X}^+}(\mathcal{S}) - \mathcal{S} \tag{4.6}$$

the shelter can be reached through the sun. In our ducklings example, a glade is formed by the water around the nest. There, the ducklings need to swim against the flow to get back inside the nest and meanwhile they stay inside the desirable region.

Another region inside the *upstream* are the *lakes*

$$\mathcal{L} = \mathcal{U} \cap \mathcal{M} - \mathcal{S} - \mathcal{G}. \tag{4.7}$$

There, the shelters are reachable through the undesirable region only. However, one can decide to stay within the desirable region forever but will always need the management options. As already described in Section 4.2, this leads to a qualitative choice, the *lake dilemma*: Inside of a lake one has to choose between eventual safety and uninterrupted desirability. Deciding for the first requires crossing the undesired region with all its consequences. But deciding for the second implies that there will always be a need for management options.

Furthermore, there are two kinds of lakes, *time-limited lakes* $\mathcal{L}_l$ and *time-unlimited lakes* $\mathcal{L}_u$

$$\mathcal{L}_u = \mathrm{Viab}_{\mathcal{U}}(\mathcal{L}), \tag{4.8}$$

$$\mathcal{L}_l = \mathcal{L} - \mathcal{L}_u. \tag{4.9}$$

Because time-unlimited lakes are the viability kernel of the lakes, there is no time pressure for a decision on the dilemma. In contrast, there is a deadline in time-limited lakes. This leads to a stronger form of the dilemma called the *pressing lake dilemma*. Within the metaphorical ducklings example in Figure 4.3 the time-unlimited lake is the waterfall starting from the glade and the subsequent calmer waters. The ducklings can swim in the calmer waters as long as they want to. The time-limited lake is the following waterfall

that splits into two streams. Here, the ducklings drop down the waterfall and they have only a moment to decide for the left or the right.

The rest of the upstream is split into the *remaining sunny* and the *dark upstream*

$$\mathcal{U}^{(+)/-} = (\mathcal{U} - \mathcal{M}) \cap \mathcal{X}^{+/-} \tag{4.10}$$

depending on whether the actual state is inside the desirable or undesirable region. Within the ducklings example, both, the remaining sunny and the dark upstream, are present.

The *Backwaters*

$$\mathcal{W} = \mathcal{M} - \mathcal{U} \tag{4.11}$$

belong to the manageable region as well as to the downstream. It is possible to stay in the desirable region by managing but the shelters are not reachable. The calm waters in the left lower corner of Figure 4.3 is a backwater because the ducklings can swim against the stream and stay inside the backwater but due to the waterfalls, they cannot reach the nest.

Analogously to the upstream, the rest of the downstream is divided into the *remaining sunny* and the *dark downstream*

$$\mathcal{D}^{(+)/-} = (\mathrm{Capt}_{\mathcal{U}}(\mathcal{W}) - \mathcal{U} - \mathcal{W}) \cap \mathcal{X}^{+/-}. \tag{4.12}$$

In Figure 4.3, the part of the waters that belongs to the undesirable region of the downstream is the dark downstream.

If the desirable region can be reached over and over again one is inside the *eddies* $\mathcal{E}$ that are divided into *sunny eddies* $\mathcal{E}^+$ and *dark eddies* $\mathcal{E}^-$. The metaphorical image behind the naming is that of a circular flow where one part is in the desirable and the other part in the undesirable region. Following this metaphorical image, we define them as the maximal pair of sets fulfilling

$$\mathcal{E}^{+/-} \subseteq \mathcal{X}^{+/-} - \mathcal{U} - \mathcal{D}, \tag{4.13a}$$

$$\mathcal{E}^{+/-} \subseteq \mathrm{Capt}_{\mathcal{U}}\left(\mathcal{E}^{-/+}\right), \tag{4.13b}$$

$$\mathcal{E} = \mathcal{E}^+ + \mathcal{E}^-. \tag{4.13c}$$

$\mathcal{E}^{+/-}$ are implicit capture basins (in the sense of Equations (2.19a) and (2.19b)). This feature will become relevant for the application of the Saint-Pierre algorithm (cf. Section 4.5.2). The worst regions are the *trenches*

$$\Theta = \mathcal{X} - \mathrm{Capt}_{\mathcal{U}}\left(\mathcal{X}^+\right) \tag{4.14}$$

because once inside one cannot reach the desirable region ever again. Inside the *abysses*

$$\Upsilon = \mathcal{X} - \mathcal{U} - \mathcal{D} - \mathcal{E} - \Theta \tag{4.15}$$

one can reach the desirable region a finite number of times only, and again it is distinguished between sunny and dark abysses $\Upsilon^{+/-} = \Upsilon \cap \mathcal{X}^{+/-}$.

The 5 sets $\mathcal{U}$, $\mathfrak{D}$, $\mathfrak{E}$, Υ, and Θ together are called the *main cascade*. The combination of all the mentioned regions in this section is called *main partition* of the TSM framework. In contrast, there exists an even finer distinction of regions with regard to mutual reachability that is elaborated in [Hei2016, pp. 28-29].

Differences to the original definition. The variant definition of TSM presented in this chapter differs from the original definition in [Hei2016], so a note on the difference is due. These are: (a) using control theory's notion of reachability instead of "safe reachability" (as defined in [Hei2016]). Also, we require a target set be reached in finite but arbitrary large time, instead of infinite time because it simplifies the computation and is more realistic. But for computationally supported estimations of specific models these differences are usually not relevant.

An extension of TSM in comparison to the original framework is the distinction of lakes in finite-time and infinite-time lakes and the subsequent introduction of the novel pressing lake dilemma.

4.4 2d-Examples

In this section, we will apply the TSM-framework to multiple illustrative examples. To show the broad applicability, the models range topic-wise from natural, co-evolutionary, socio-economic to mechanical. These models have all been chosen to be 2-dimensional in order to ensure a 2-dimensional plot of the state space can be done. This low-dimensionality implies that the models do not focus on realism but are just meant to give the reader an impression of the possible partitions that can occur. Furthermore, the analysis was done manually.

4.4.1 Layout of the Phase Portraits

In the following, the plots are structurally different, but their main layout is the same to ease the understandability. The default flow is shown as a light blue, thick streams while the management flows are shown as thin, dark blue, dashed/dotted streams. The stable fixed points of each flow are marked with round points in the corresponding color. The regions are colored with the color scheme introduced in Figure 4.2.

4.4.2 Global Carbon Cycle

As a first example, we start with the global carbon cycle that has already been introduced in Model Description 2 on page 26. The model is placed in the natural sciences and describes the carbon cycle with its three main stocks, the terrestrial $c_t = c_{terrestrial}$, marine $c_m = c_{marine}$ and the atmospheric carbon $c_a = c_{atmospheric}$, with their complicated nonlinear interactions on a global level. As we use the pre-industrialized version, the sum of all carbon stocks remains constant $C = c_t + c_m + c_a = const$, hence we can reduce the number of dynamic variables by taking c_a as a dependent variable implicitly defined by

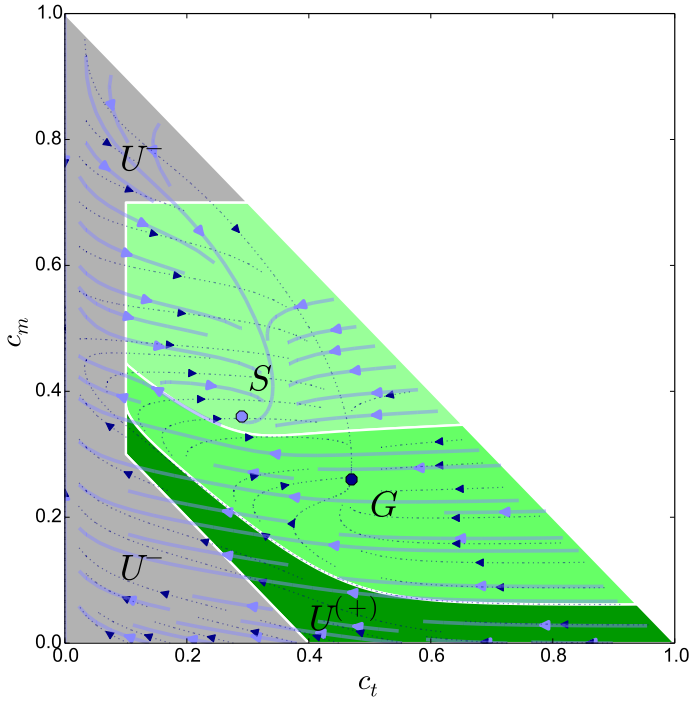


Figure 4.4: TSM-Partition of the model "Global Carbon Cycle" as explained in Section 4.4.2. All of the state space is part of the upstream. In detail, it is composed of a shelter, a glade, a dark upstream and a sunny remaining upstream. As explained in Section 4.4.1, the light-blue stream corresponds to the default dynamics with the light-blue, stable fixed point and the dark-blue stream to the management dynamics with the dark-blue, stable fixed point. The coloring and naming of the regions follows Figure 4.2. This graphic was published in [Hei2016].

c_t and c_m. Due to the complexity of the model, we recall only the two major dynamical equations here:

$$\dot{c}_t = NEP(p, r, c_t) - \alpha c_t, \tag{4.16a}$$
$$\dot{c}_m = I(c_a, c_m). \tag{4.16b}$$

Find the detailed equations in Model Description 2 on page 26 and the parametrization in Table 3.1 on page 27.

To enable the analysis of TSM, we need to distinguish between default dynamics and management. The only parameter with human influence in the model is the harvesting parameter α, so we set the default dynamics to $\alpha^{\mathrm{d}} = 0.5$ and $\alpha^{\mathrm{m}} = 0.25$. Hence management means here that humans cannot harvest as much terrestrial carbon, e.g. through agriculture.

Further, we need to introduce a desirable region. Within this model we use three

Parameter	Fig. 4.5 (a)	Fig. 4.5 (b)	Fig. 4.5 (c)	Fig. 4.5 (d)
α	-0.1	-0.15	-0.15	-0.15
ϕ	4	4	4	4
κ	12000	6000	6000	6000
r	0.04	0.04	0.04	0.04
γ^{d}	$4 \cdot 10{-}6$	$8 \cdot 10{-}6$	$8 \cdot 10{-}6$	$8 \cdot 10{-}6$
γ^{m}	$2.8 \cdot 10{-}6$	$13.6 \cdot 10{-}6$	$16 \cdot 10{-}6$	$11.2 \cdot 10{-}6$
x_{min}	1000	1200	4000	4000
y_{min}	3000	2000	3000	3000

Table 4.1: The parameters of the of the model "Combined Population and Resource Dynamics" used for the TSM analysis (cf. Model Description 6). As there were 4 different analysis with 4 different sets of parameters done, they are shown for each of Figures 4.5 (a) to 4.5 (d) separately.

borders motivated by the PBs: $c_a < 0.6$, $c_m > 0.7$ and $c_t > 0.1$. This implies we demand there to be enough terrestrial carbon, e.g. plants, but not too much carbon dioxide pollution in the air or the oceans.

We remind the reader that due to the low complexity of the models, they do not provide a realistic dynamics but their idea is more conceptual and illustrative. The same holds for the choice of the aforementioned parameters.

In Figure 4.4, the different flows and the resulting regions are depicted. We refer the reader to the explanation on how to read these figures in Section 4.4.1.

In this example, the default dynamics has a stable fixed-point (marked by a light-blue dot) that is in the desirable region. For that reason, all the points from where one can reach this fixed point without transgressing the boundary and using only the default dynamics (light-blue flow) are the shelter.

The management dynamics has only one attractor: a stable fixed point (dark-blue dot) within the desirable region. From there, by first following the default dynamics and then switch shortly before reaching the undesirable region, it is possible to (a) reach the shelter and (b) without going into the undesirable region. (a) implies that the whole state space is part of the upstream as I can always reach the shelter. And (b) implies that the region containing the management fixed point is the glade.

In the other regions, either one is in the undesirable region or one has to cross it to reach the shelter, so they are classified as sunny remaining and dark upstream. There is no lake dilemma in this example.

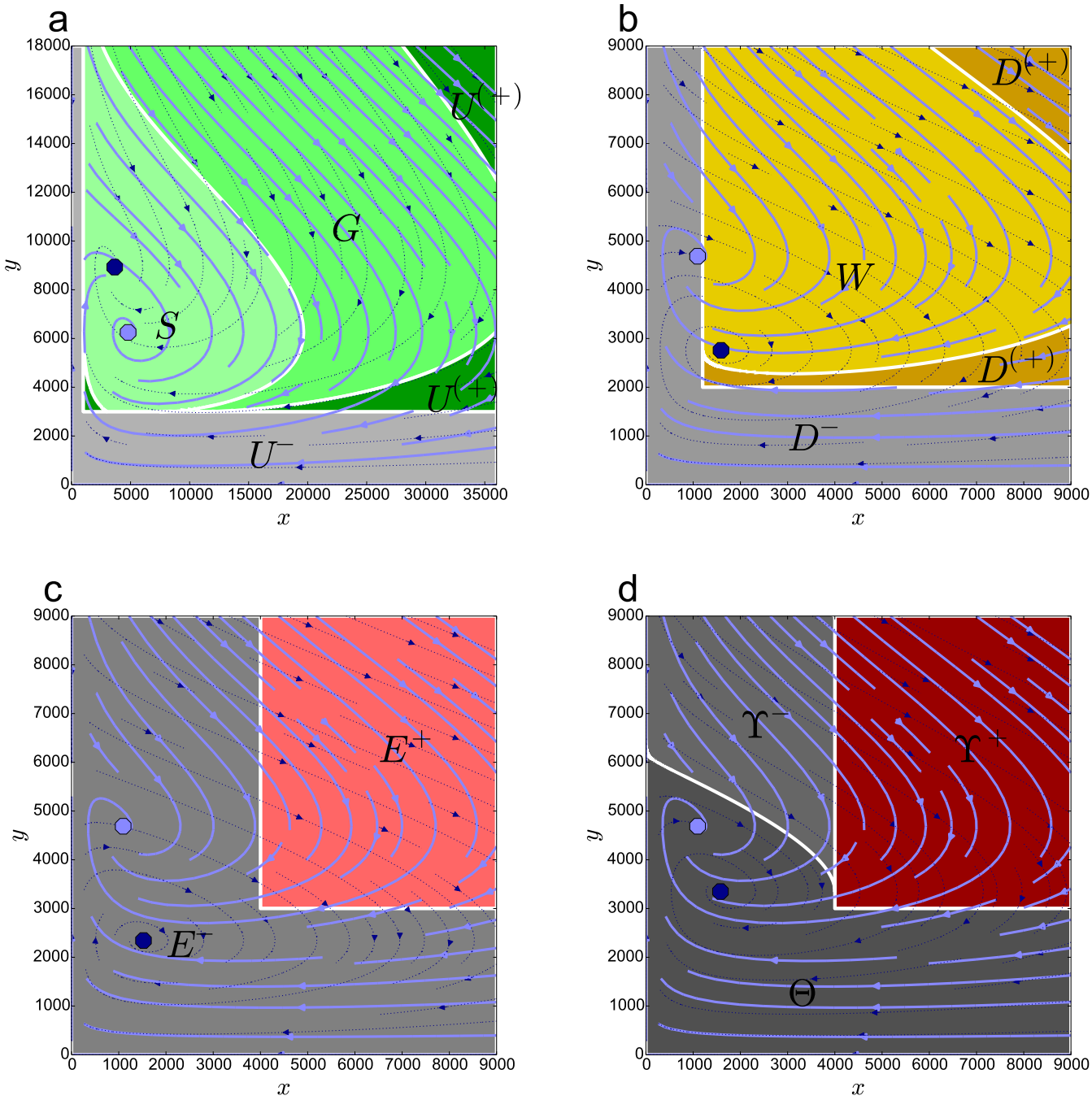


Figure 4.5: TSM-partition of four different parameterizations of the model "Combined Population and Resource Dynamics" (cf. Model Description 6). The different parameters are summarized in Table 4.1. In (a) the whole phase space belongs to the upstream, in (b) to the downstream, in (c) to the eddies and in (d) to the abysses and trenchs. While all dynamics do not differ qualitatively, the quantitative difference due to the changing parameters shows a striking impact on the resulting TSM-partition. As explained in Section 4.4.1, the light-blue stream corresponds to the default dynamics with the light-blue, stable fixed point and the dark-blue stream to the management dynamics with the dark-blue, stable fixed point. The coloring and naming of the regions follows Figure 4.2. These graphics were published in [Hei2016].

4.4.3 Combined Population and Resource Dynamics

The second example is a co-evolutionary model, i.e. it combines natural and socio-economic parts. Please find the details in Model Description 6.

Model Description 6 (Combined Population and Resource Dynamics)**.** This model was developed in [Bra1988] to explain the rise and fall of the native civilization on Rapa Nui, Easter Island, before western contact. It is derived from simple economic principles and leads to a Lotka-Volterra-type predator prey model with a finite resource. The human population x is preying on the island's forest stock y which is growing logistically

$$\dot{x} = \alpha x + \phi\gamma xy \tag{4.17}$$

$$\dot{y} = ry(1 - y/\kappa) - \gamma xy \tag{4.18}$$

where α is the difference between the natural reproduction and death rate (negative) of the human population (without food), γ the parameter how strongly the humans prey, ϕ the efficiency on how the food from preying increases the human population, r the difference between the natural reproduction and death rate (positive) of the forest and finally κ the maximal forest cover of the island. The time scales of that model was chosen to be 10 years, so an evolution over $t = 2$ means over 20 years [Bra1988]. The parametrization used within the analysis is summarized in Table 4.1. The parameters of the default flow in Figure 4.5 (a) correspond to the original parameters chosen in [Bra1988].

As the model is rather simple (cf. dynamics in Figures 4.5 (a) to 4.5 (d)), for each of the different choices of parameters the dynamics actually does not change qualitatively. In particular, for each dynamics, there exists one stable fixed point that is the only attractor.

Still, with these minor changes, we managed to find 4 completely different partitions. The different parameterizations have been chosen to demonstrate interesting results and not for the sake of realism. In Figure 4.5 (a), the whole state space is part of the upstream and no lake dilemma found. In Figure 4.5 (b), the whole state space is part of the downstream as the light-blue fixed-point of the default dynamics is just in the undesirable region but the dark-blue fixed point of the management dynamics is within the desirable region. Hence, by keeping up the management one can stay inside the desirable region forever. In Figure 4.5 (c), the whole space is composed of eddies. The reason is that from the default fixed point one can reach the desirable region shortly by switching to the management dynamics. And from the management fixed point, the default fixed point is reachable again. Still, not even by combination of the two dynamics, it is possible to stay within the desirable region forever, so everything is classified as eddies. The state space of Figure 4.5 (d) has three parts: the trench, where it is impossible to leave the undesirable region, the sunny abyss, where one is in the sun but will enter the trench finally, and the dark abyss, where one can reach the sunny abyss but from there

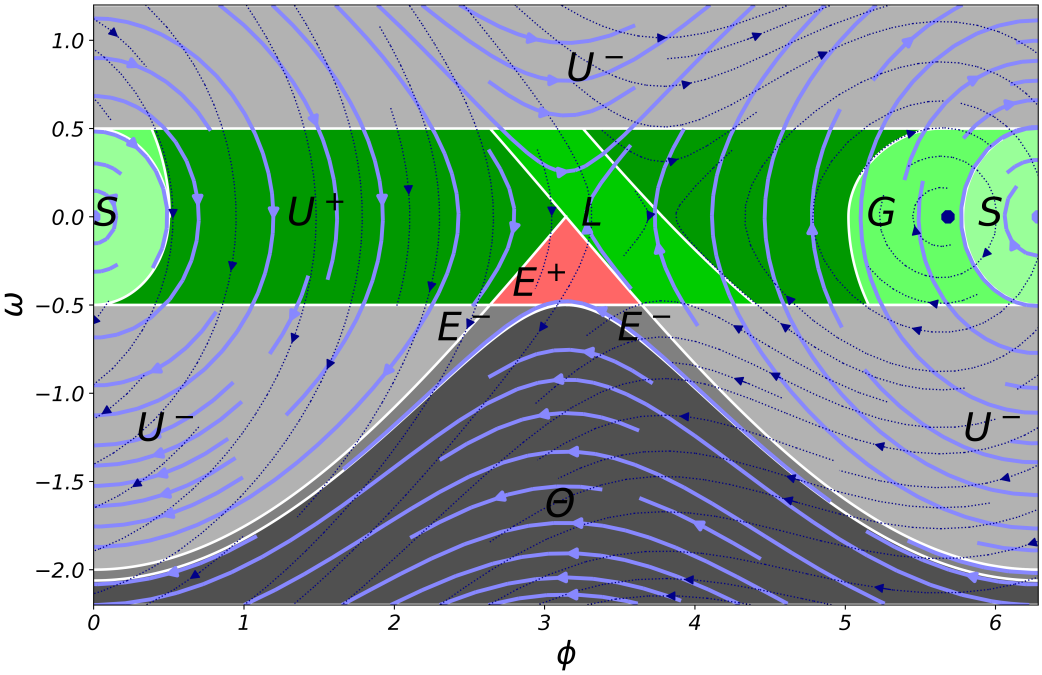


Figure 4.6: The Analysis of the model "Gravity Pendulum" (cf. Model Description 7) with respect to Topology of Sustainable Management. In particular, the states space exhibits a lake, giving rise to the lake dilemma. Further, there are eddies, a region where one may reach the sunny part over and over again, but will always have to go in the dark part, too. As explained in Section 4.4.1, the light-blue stream corresponds to the default dynamics with the light-blue, stable fixed point and the dark-blue stream to the management dynamics with the dark-blue, stable fixed point. The coloring and naming of the regions follows Figure 4.2. This graphic was published in [Hei2016].

one will enter the trench finally.

4.4.4 Gravity Pendulum

The third illustrative example is the gravity pendulum taken from mechanical physics. Details are in Model Description 7.

Model Description 7 (Gravity Pendulum)**.** This model is a undamped pendulum with fixed length 1, maxx 1 and the gravity constant as 1, too. Further we assume a driving with parameter a, which is set to $a^{\mathrm{d}} = 0$ for the default dynamics and

$a^{\mathrm{m}} = 0.6$. Hence, the dynamics is given by

$$\begin{aligned} \dot{\phi} &= \omega, \\ \dot{\omega} &= -\sin\phi - a. \end{aligned} \tag{4.19}$$

These equations already show, that this is formally equivalent to the "Single Generator in a Power Grid" (cf. Model Description 3 on page 28) without damping.

Further, we assume the undesirable region to be given by a too strong angular speed ω of the pendulum, hence we take as condition for the desirable region $|\omega| < \omega_{max} = 0.5$.

This model, whose state space is depicted in Figure 4.6, shows a rich assortment of regions. The shelter is when the pendulum is close to or at the lower position of the pendulum ($\phi = 0$ or equivalently $\phi = 2\pi$). When being further away of the shelter in the negative direction of ϕ, one can use the driving (management) to brake slowly into the shelter. Further, there is a lake next to the fixed point of the default dynamics at $(\phi, \omega)^\top = (\pi, 0)^\top$ (top of the pendulum), where one can stay by repeatedly falling down a bit (default) and then pushing up again with the driving (management).

Finally, there are eddies. In these regions, just by using the default dynamics, one can reach the desirable region over and over again, but will have to enter the undesirable region in-between. Note that using the management would push the trajectory in the trench. This situation basically corresponds to a rather quickly turning pendulum in direction of the driving. When the pendulum is approaching the top, it slows down as some of the energy is converted to potential energy, so the absolute value of the angular speed $|\omega|$ is reduced below ω_{max} and the trajectory is temporarily inside the desirable region. After passing the top, the pendulum accumulates angular speed again and $|\omega|$ exceeds ω_{max}, hence leaving the desirable region. As the turning is in direction of the driving, the usage of management increases the angular speed and may kick the system into the trench.

The trench corresponds to a similar situation as the eddies. The difference is that there is so much kinetic energy that the (absolute value of the) pendulums angular speed does not drop below ω_{max} even at the top. Hence, the trajectory stays in the undesirable region forever.

The lake implies the existence of a lake dilemma. Here, this is the qualitative choice between risking temporarily the high angular speed but finally ending up in the shelter and needing management for the procedure of falling down and pushing up with driving (as described before).

4.4.5 Substitution of a Dirty Technology

The next example is an economic model between two goods, e.g. electric energy, where one is "dirty" and the other "clean". Also, there is an political level by allowing subsidies for the "clean" technology. Details are in Model Description 8.

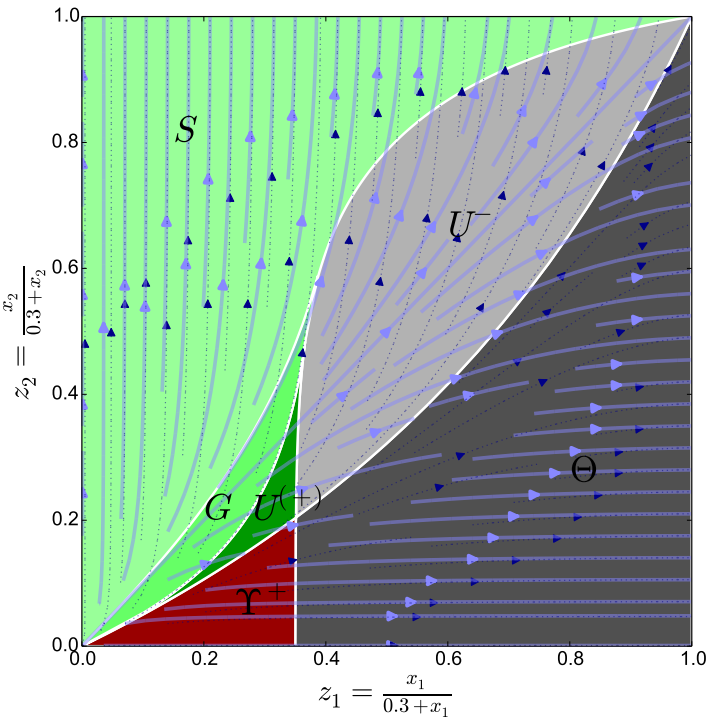


Figure 4.7: The resulting TSM partition of the model "Substitution of a Dirty Technology" (cf. Model Description 8). As the state space is unbounded with $x \in \mathbb{R}^2_{\geq 0}$, the dynamics has been rescaled such that $z_i \to 1$ represents $x_i \to \infty$. This is the same transformation as the proposed compactification in Section 4.5.3. In particular, the sunny abyss Υ^+ corresponds to an economic lock-in effect where even with subsidies (management) for the clean technology 2, the dirty technology 1 will have the dominant market position. As explained in Section 4.4.1, the light-blue stream corresponds to the default dynamics and the dark-blue stream to the management dynamics. The coloring and naming of the regions follows Figure 4.2. This graphic was published in [Hei2016].

Model Description 8 (Substitution of a Dirty Technology)**.** This model is purely socio-economic. Two goods, e.g. electric energy, that are perfectly substitutable, are competing in the same market. The technologies for the production of one is assumed to be "clean" while the other is "dirty". The cost functions C_1 and C_2 are convex functions of the production output y_i and decrease with a high cumulative production x_i (meaning $\dot{x}_i = y_i$) via learning-by-doing

$$C_i(y_i) = \gamma_i \frac{y_i^{1+\sigma_i}}{(1+\sigma_i)x_i^{\alpha_i}}, \tag{4.20}$$

where γ_i are cost factors, σ_i the convexity parameters and α_i learning exponents.

Following standard economic assumptions, we consider that demand equals production $D = y_1 + y_2$ ("market clearance"), that prices are equal $p = p_1 = p_2$, that demand depends linearly on the price $D(p) = D_0 - ap$ and that the price equals the marginal cost $p_i = \frac{\partial C_i}{\partial y_i} = \gamma_i \frac{y_i^{\sigma_i}}{x_i^{\alpha_i}}$.

One can then uniquely solve for the produced amounts $y_i = f_i(x_1, x_2)$. Plugging in $\dot{x}_i = y_i$, leads then to a two-dimensional system of ordinary differential equations

$$\dot{x}_i = f_i(x_1, x_2) \tag{4.21}$$

for $i = 1,\ 2$.

The parameters are chosen as follows: $D_0 = 1$, $a = 1$, $\sigma_i = 0.2$, $\alpha_i = 0.5$ and $\gamma_1 = 1$. Further, for the default dynamics we take ${\gamma_2}^{\mathrm{d}} = 1$.

Until now, the model is symmetric under the exchange of 1 and 2. Now, let us consider technology 1 to be the dirty one and 2 to be the clean one. As management option, we consider lowering γ_2 to ${\gamma_2}^{\mathrm{m}} = 0.5$, representing subsidies in order to induce technological change [Jaf2002; Kal2012].

In the desirable region we encode the goal of keeping the usage of the dirty technology below some limit, i.e. $y_1 < y_{1,max}$. This constraint can be reformulated in terms of the cumulative productions x_i as

$$x_2 > \begin{cases} x_1 \left(\frac{1-y_{1,max}}{y_{1,max}} - \frac{1}{y_{1,max}^{\frac{4}{5}} \sqrt{x_1}} \right)^{\frac{2}{5}} & \text{if } x_1 > \frac{y_{1,max}^{\frac{2}{5}}}{(1-y_{1,max})^2} \\ 0 & \text{otherwise.} \end{cases} \tag{4.22}$$

The state space of the model "Substitution of a Dirty Technology" is shown in Figure 4.7. Note that due to the transformation in the picture, the upper border represents $x_1 \to \infty$ and the right border $x_2 \to \infty$. Except for the upper right corner, reaching the top border represents the (relative) die-out of the technology 1, as this one goes only to a finite value. Vice versa, reaching the right border corresponds to a (relative) die-out of technology 1.

Thus, we observe a shelter in the upper part with a neighboring glade. Furthermore, the system exhibits an abyss and a trench. These correspond to a so-called lock-in effect, when the dirty technology is so strongly established (relatively to the clean one) that even continuous subsidies cannot ensure the dominance the clean technology.

Note that the abyss actually reaches the origin. That is due to the system not being developed for the starting situation of such a technology but when both technologies are a bit established, but one more than the other

4.4.6 Competing Plant Types

Finally, the 5$^{\text{th}}$ and last two-dimensional example is from ecology. It concerns two fictitious, competing plant types that change the soil to their own benefit and are harvested.

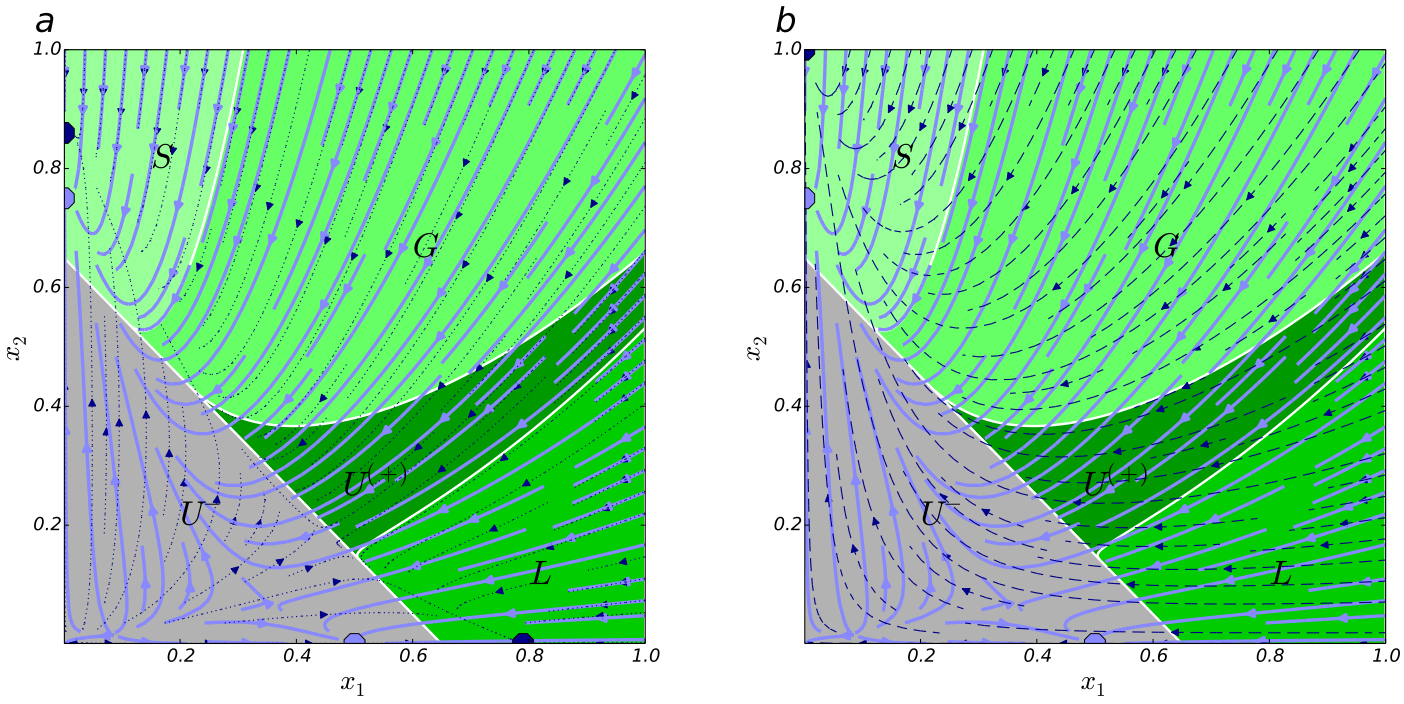


Figure 4.8: The TSM-Analysis of the model "Competing Plant Types" (cf. Model Description 9). The lake in the lower right corner is induced by the two different management options and the r parameter, describing the asymmetry of the plants productivities. When in the lake, the harvesters can decide to either have continuously reduced, but sufficient harvest, the first management, or to protect plant type 2, leave it to grow and only harvest plant type 1. But then, they will temporarily have a reduced and insufficient harvest, meaning they cross the dark region. Still, they will finally end up in the shelter, where they can harvest plant type 2 without restriction. As explained in Section 4.4.1, the light-blue stream corresponds to the default dynamics with the light-blue, stable fixed point and the dark-blue stream to the management dynamics with the dark-blue, stable fixed point. The coloring and naming of the regions follows Figure 4.2. These graphics were published in [Hei2016].

Details are in Model Description 9.

Model Description 9 (Competing Plant Types)**.** The two plant types 1 and 2 in this model are fictitious. They compete for the soil by modifying it to their own benefit and the others disadvantage. Their growth is following a logistic-type dynamics with land-cover portions x_1 and x_2

$$\dot{x}_1 = x_1 \left(K_1 \left(x_1, x_2\right) - x_1\right) - h_1 x_1 \tag{4.23a}$$
$$\dot{x}_2 = r x_2 \left(K_2 \left(x_1, x_2\right) - x_2\right) - h_2 x_2, \tag{4.23b}$$

where r is a productivity constant, $h_{1/2}$ the harvest rates and the two dynamic capacities

$$K_1 \left(x_1, x_2\right) = \sqrt{x_1} \left(1 - x_2\right) \tag{4.24a}$$
$$K_2 \left(x_1, x_2\right) = \sqrt{x_2} \left(1 - x_1\right). \tag{4.24b}$$

The latter represent the fact of modifying the soil to ones own benefit. $r = 2$ is chosen to represent that plant type 2 is more productive, i.e. grows relatively quicker than plant type 1.

In this model, the management is done via the harvesting of the plants. In the default dynamics, we set ${h_1}^{\mathrm{d}} = {h_2}^{\mathrm{d}} = 0.2$. The first management is to halve both values to ${h_1}^{\mathrm{m}} = {h_2}^{\mathrm{m}} = 0.1$, representing a stricter harvesting policy to aim for higher harvests in the long-term. The second management option is ${h_1}^{\mathrm{m}'} = 0.4$ and ${h_2}^{\mathrm{m}'} = 0$, i.e. a protection of plant type 2 in order to allow plant 2 to grow more and change the plant mix.

Finally, the desirable region is set to be at $x_1 + x_2 > x_{min} = 0.65$ to ensure some minimum harvest.

In this model, we consider more than one management. For that reason, we decided to split up the state space plot in Figure 4.8. In Figure 4.8 (a) the default flow and the first management is shown, while in Figure 4.8 (a) the default flow and the second management is shown. In both cases, the resulting TSM-partition is depicted. Note, that both state space plots show the same state space, just with differing management options drawn in. But at any time, both management options are available.

A shelter exists due to the higher productivity of plant type 2, and the whole state space is part of the upstream because of the second management option.

The interesting feature in this system is the natural existence of a lake when having a relatively large amount of type 1, giving rise to the lake dilemma. In this case, the harvester can decide whether they harvest little (with the first management option), but barely enough to stay in the desirable region. Or adjust their harvest preference by protecting plant type 2 (with the second management option). But that implies they have to go temporarily through the undesirable region as they can only harvest plant type 1

and plant type 2 needs a while to adjust the soil to its own benefit. So the harvesters will temporarily have not enough harvest available, but finally when reaching the shelter there will be no need for restriction on the harvest anymore.

4.5 Saint-Pierre Algorithm and Extension

In order to take the step from the manually analyzed 2-dimensional examples to a three dimensional model, an automation of the analysis is due. For that reason we introduced TSM with the variant definition based on VT's Capture Basin and Viability Kernel in Section 4.3. This enables us to apply the Saint-Pierre algorithm (SPA) that was developed for the estimation of these two basic concepts of VT [Sai1994]. In this section, we shortly will review the algorithm and then present extensions that we will need for the analysis of the 3-dimensional model afterwards.

4.5.1 Sketch of the Algorithm

Again, we work with a general control system (Equation (2.15), here only repeated)

$$\dot{\boldsymbol{x}} = h\left(\boldsymbol{x}, \boldsymbol{u}\right). \tag{4.25}$$

The algorithm starts with a finite discretization $\mathcal{Y}^r$ of the constraint set $\mathcal{Y}$ where a point $\boldsymbol{x} \in \mathcal{Y}$ is at most at a distance r of a point $\boldsymbol{y} \in \mathcal{Y}^r$ *. Furthermore, a small time step $\Delta t > 0$ is chosen and it supposes that the set of controls $\mathcal{U}$ is discrete (if not, it is also discretized). It supposes that h is l-Lipschitz and there exists an upper bound M of $\mathcal{Y}$.

The algorithm starts by computing, for each point $\boldsymbol{x} \in \mathcal{Y}^r$ and for each control $\boldsymbol{u} \in \mathcal{U}$, the successors $S(\boldsymbol{x}, \boldsymbol{u})$ of $\boldsymbol{x}$ when applying control $\boldsymbol{u}$, for a linearized, extended dynamics defined from h after small time step Δt. It is extended in the sense that the successors include all the points located in a ball around $\boldsymbol{x} + h(\boldsymbol{x}, \boldsymbol{u}) \cdot \Delta t$, hence

$$S(\boldsymbol{x}, \boldsymbol{u}) = \left\{ \boldsymbol{y} \in \mathcal{Y}^r \mid \|\boldsymbol{y} - (\boldsymbol{x} + h(\boldsymbol{x}, \boldsymbol{u}) \cdot \Delta t)\| \leq r + \frac{Ml}{2} \cdot (\Delta t)^2 \right\}. \tag{4.26}$$

This extension of the dynamics guarantees that the algorithm described below converges to the actual viability kernel when Δt and the resolution of the grid decrease to 0 [Sai1994]. Computing and storing all the successors for each point of the grid rapidly becomes computationally heavy when the dimensionality of the state space is large and the grid resolution is small†.

Then, the algorithm builds a series of discrete sets (subsets of the grid $\mathcal{Y}^r$) $Q_0 = \mathcal{Y}^{\mathrm{h}}$, Q_1, $Q_2, \ldots$ such that $Q_{i+1} \subset Q_i$, defined as follows:

$$Q_{i+1} = \{\boldsymbol{x} \in Q_i \mid \exists \boldsymbol{u} \in \mathcal{U} \colon S(\boldsymbol{x}, \boldsymbol{u}) \cap Q_i \neq \emptyset\} \tag{4.27}$$

*Often $\mathcal{Y}^r$ is chosen to be a regular grid of mesh-width r.

†This is often referred to as "curse of dimensionality".

After a finite number of steps, the algorithm reaches a fixed point, i.e. $Q_{i+1} = Q_i$. The set Q_i is the viability kernel of $\mathcal{Y}^r$ for the linearized, extended discrete dynamics. In [Sai1994], it is shown, that this set converges to the viability kernel of the continuous time dynamics when Δt and r tend to 0 appropriately. Note that the approximations are done from the exterior of the viability kernel: generally, the approximation includes points that do not belong to the actual viability kernel but their proportion decreases when the resolution of the grid increases.

This algorithm has been extended in [Def2007] for using continuous sets Q_i, using a machine learning algorithm that takes as input the points of the grid that belong to Q_i and the ones that do not, and derives an approximation of its boundary. This opens the possibility to represent continuous viability kernels that are defined more conveniently than a huge set of points.

A slight modification of the of the algorithm described above enables us to approximate the capture basin. Again, we start with a finite discretization of the constraint set $\mathcal{Y}^r$ and define the discretized target set $\mathcal{X}^r = \mathcal{X} \cap \mathcal{Y}^r$ for a target set $\mathcal{X}$. Again, we create a series of discrete sets Q'_i with $Q'_0 = \mathcal{X}^r$ and where the successors of all elements in Q'_i are in Q'_{i+1}

$$Q'_{i+1} = \left\{ x \in \mathcal{Y}^r \mid \exists u \in \mathcal{U} \colon S(x,u) \cap Q'_i \neq \emptyset \right\}. \tag{4.28}$$

Again, after a finite number of steps, the algorithm reaches a fixed point, i.e. $Q'_{i+1} = Q'_i$, and Q'_i is the capture basin of $\mathcal{X}^r$ with the constraint set $\mathcal{Y}^r$ for the linearized extended discrete dynamics. In contrast to the viability estimation, this is an approximation from the interior.

Improvements and extensions to this algorithm are currently under intensive research. Relations to dynamical programming [Fra1989] and other extensions [Bay2002; Sai2001] can provide the minimal time to reach the target set. Also, one can even find controllers that drive the system to the target set [Car1998; Cha2011; Lho2007].

4.5.2 Estimation of Implicitly Defined Capture Basins: Eddies

Eddies are defined as implicit capture basins in Equations (4.13a) and (4.13b) (in the sense of Equations (2.19a) and (2.19b)). The SPA assume that the target set does not change when computing the capture basin. But as $\mathcal{E}^+$ is the target set of $\mathcal{E}^-$ and vice versa, we need to extend the SPA. In order to do that, we develop an alternate definition in terms of a limit process.

We start by defining the largest sets that could contain eddies

$$\mathcal{E}^+{}_0 = \mathcal{X}^+ - \mathcal{U} - \mathcal{D}, \tag{4.29a}$$
$$\mathcal{E}^-{}_0 = \mathcal{X}^- - \mathcal{U} - \mathcal{D}, \tag{4.29b}$$

and then use the iteration step

$$\mathcal{E}^-{}_i = \mathrm{Capt}_{\mathcal{U}(}\mathcal{E}^+{}_{i-1}) \cap \mathcal{E}^-{}_{i-1}, \tag{4.30a}$$
$$\mathcal{E}^+{}_i = \mathrm{Capt}_{\mathcal{U}(}\mathcal{E}^-{}_i) \cap \mathcal{E}^+{}_{i-1} \tag{4.30b}$$

for $i = 1, 2, \ldots$. Note that ${\mathscr{C}^-}_i$ used in Equation (4.30b) is already computed in Equation (4.30a). So this can really be seen as a step-by-step prescription. Thus, the eddies can be recovered as

$$\mathscr{C}^+ = \lim_{i \to \infty} {\mathscr{C}^+}_i, \tag{4.31a}$$

$$\mathscr{C}^- = \lim_{i \to \infty} {\mathscr{C}^-}_i. \tag{4.31b}$$

The limit exists because both sequences are monotone and non-increasing.

This limit process can be interpreted as an imperative description of subsequently applying the SPA and used as an algorithmic description for the estimation of implicit capture basins. The SPA works on a discretized state space with finitely many elements. Hence, the existence of the limit provides that there exists an $k \in \mathbb{N}$ such that $\mathscr{C}^+ = {\mathscr{C}^+}_k = {\mathscr{C}^+}_{k-1}$ and $\mathscr{C}^- = {\mathscr{C}^-}_k = {\mathscr{C}^-}_{k-1}$ and the algorithm converges after a finite number of steps.

This iteration process follows the idea of being able to visit the sunny part over and over again and is an algorithmic description for the estimation of eddies. Other similarly implicitly given sets can be estimated by adjusting this basic idea.

4.5.3 Compactification

In Section 4.5.1, a finite discretization of the constraint set was assumed. As the constraint set might be unbounded, this could pose a problem. In this section, we will propose a compactification of the state space to solve this problem. It will be used for the 3-dimensional example in Section 4.6.

There are multiple ways to map an unbounded state space to a bounded one, depending on the specific need for the system. In case of the example system analyzed later, each coordinate is bounded from below and unbounded from above, so we will focus on that. This is rather common in socio-economic models, in particular due to continuous economic growth. We propose a solution that maps each coordinate separately.

The compactification can be done for each control $\boldsymbol{u} \in \mathcal{U}$ separately, so in this section, we will work on a general dynamic system. Hence, for a fixed control $\boldsymbol{u}^{\mathrm{c}} \in \mathcal{U}$, we take $f(\boldsymbol{x}) = h(\boldsymbol{x}, \boldsymbol{u}^{\mathrm{c}})$ and reduced the control system (Equation (2.15)) to an ordinary differential equation (Equation (2.6), here only repeated)

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}) \tag{4.32}$$

with $\boldsymbol{x} \in \mathfrak{X}$. In particular, we demand $\mathfrak{X} = \mathbb{R}^n_{\geq 0}$ as we just mentioned that we want the system to be bounded from below and unbounded from above. Then we propose the coordinate transformation

$$\begin{aligned} \Phi : \mathbb{R}^n_{\geq 0} &\longrightarrow [0,1)^n \\ (x_i) &\longmapsto \left(\frac{x_i}{x_i^{\mathrm{mid}} + x_i}\right), \end{aligned} \tag{4.33}$$

where $x_i^{\text{mid}} \in \mathbb{R}_{\geq 0}$ are parameters. Applying this transformation on the dynamics leads to a new set of ordinary differential equations

$$\begin{aligned} \dot{y} &= \tilde{h}(y) := ((D\Phi \cdot f) \circ \Phi^{-1})(y) \\ \implies \quad \dot{y}_i &= \frac{(1-y_i)^2}{x_i^{\text{mid}}} f\left(\frac{y_i x_i^{\text{mid}}}{1-y_i}\right) \quad \forall i \in \{0, \dots, n-1\} \end{aligned} \tag{4.34}$$

with $y \in [0,1)^n$, i.e. inside a bounded space, where $D\Phi$ is the Jacobian of Φ and $\circ$ is the symbol for function composition. The parameters x_i^{mid}, summarized to the vector x^{mid}, are precisely the scales for each coordinate that is "resolved best" because $\Phi(x^{\text{mid}}) = (\frac{1}{2}, \frac{1}{2}, \dots)^\top$. So they should be taken to be around the main region of interest.

The new system Equation (4.34) is now bounded and the application of the SPA generally possible. Still, there might be an additional problem that is introduced and solved in the following Section 4.5.4.

4.5.4 Local, Nonlinear Time Homogenization

A problem during the estimation of viability kernels from Section 4.5 is a possibly inhomogeneous time scale, i.e. that the (norm of the) RHS function of the control system (2.15) can have values through *several orders of magnitude*. For instance, models like Equation (4.34) often lead to divergences at the upper boundary of a coordinate.

This problem can actually be addressed by rescaling the time of the system in a nonlinear way. The used definitions of Viability Theory depend only qualitatively but not quantitatively on time. Hence viability kernels and capture basins in the rescaled system are equivalent to the original ones. Again, the control parameter is not necessary for the rescaling of the system. So we continue with the ordinary differential equation (Equation (4.34), here only repeated)

$$\dot{y} = \tilde{h}(y) \tag{4.35}$$

We propose to use the new system

$$\dot{y} = \tilde{\tilde{h}}(y) := \frac{\tilde{h}(y)}{\|\tilde{h}(y)\| + \epsilon}. \tag{4.36}$$

Assuming ϵ is small enough, this new system generally fulfills three criteria:

1. The systems (4.34) and (4.36) are *orbitally equivalent* (cf. [Kuz1998, Definition 2.4]), i.e. the trajectories of solutions with the same initial conditions follow the same path. In other words, only the time has been rescaled.
2. Everywhere away from fixed points, $\|\tilde{h}(y)\| \gg \epsilon$ holds and hence the time scale is properly homogenized

$$\|\tilde{\tilde{h}}(y)\| = \frac{\|\tilde{h}(y)\|}{\|\tilde{h}(y)\| + \epsilon} \approx 1. \tag{4.37}$$

3. At fixed points of the original system, the function goes to zero with the same properties as $\tilde{h}$ at that point (e.g. μ-Lipschitz or C^∞, same Lyapunov-Exponents etc.) because within a small enough environment of the fixed point $\|\tilde{h}(\boldsymbol{y})\| \ll \epsilon$ holds, thus

$$\tilde{\tilde{h}}(\boldsymbol{y}) = \frac{\tilde{h}(\boldsymbol{y})}{\|\tilde{h}(\boldsymbol{y})\| + \epsilon} \approx \frac{1}{\epsilon}\tilde{h}(\boldsymbol{y}). \tag{4.38}$$

Remark. Because the units of the coordinates of $\boldsymbol{y}$ might be different from each other, there is no real physical interpretation of $\tilde{\tilde{h}}$. But that is not necessary either as it is only an auxiliary system for the estimation with the SPA.

4.6 3d-Example: the AYS-Model

In contrast to the examples in Section 4.4, a 3-dimensional model cannot be analyzed manually anymore. Hence, we introduced TSM in terms of Viability Theory, enabling us to operationalize the TSM-framework by application of the SPA. The model's complexity has been held low so it can be used as an appropriate first example where the TSM partition is estimated automatically.

Following the title of this dissertation, the example model will be co-evolutionary. That means, it couples environmental and socio-economic dynamics [Sch1998], also referred to as the *ecosphere* and the *anthroposphere* respectively [Kuh2010]. There is a variety of co-evolutionary low-complexity models published [Bar2016; Bra1988; Kel2011; Lad2015; Nit2017; Wie2015] and we focus on climate change, economic output and energy transformation. These topics are represented in the model's variables by an atmospheric carbon stock and the latter by the (global) economic output and a knowledge stock on renewable energy production, possibly leading to a technological change. Two common climate policies, i.e. managements, can be implemented in a conceptual manner: (i) *low growth* and (ii) climate mitigation by inducing an *energy transformation*.

In such a model, there are two naturally arising boundaries. The first one is the *planetary boundary on climate change* (PB-CC) [Roc2009; Ste2015b] which limits the amount of atmospheric carbon. The second one is a wealth-related social foundation prescribing a threshold below that the yearly economic output should not fall (SF-Y).

4.6.1 Model Description

To develop such a low-complexity model of climate change, economic output and energy transformation, we took inspiration from [Kel2011] and added a renewable energy sector with a learning-by-doing dynamics. Its structure is depicted in Figure 4.9.

The first is the *excess atmospheric carbon stock* A [GtC = giga tons of carbon], measured w.r.t. a pre-industrial level $A^0 \approx 600\,\text{GtC}$. It increases with current CO_2 emissions E [GtC/a = GtC per year]. Taking A^0 as an estimate for the long-term no-emissions equilibrium value, we assume A approaches zero if $E = 0$, due to carbon uptake by

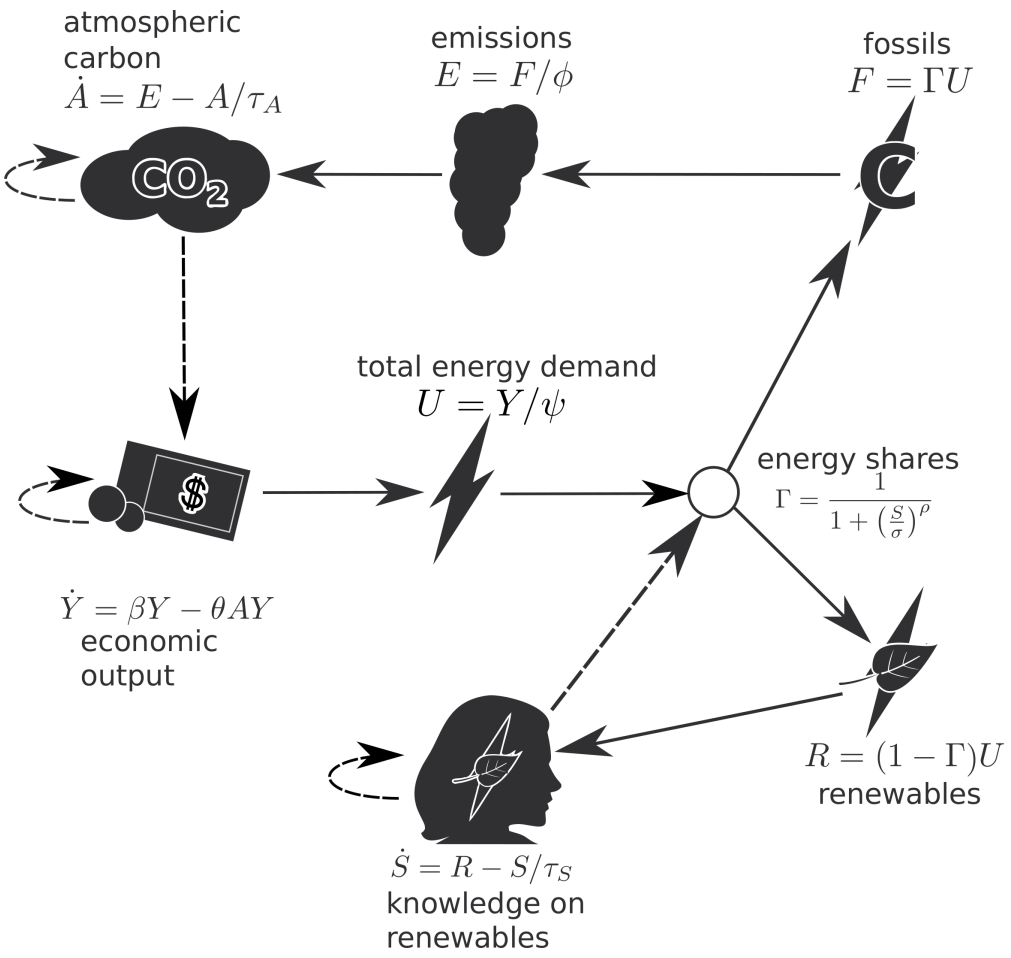


Figure 4.9: The interplay of the three dynamical variables *excess atmospheric carbon stock* A, *economic production* Y and *renewable energy knowledge stock* S and the five dependent variables *energy demand* U, *fossil (or renewable) energy flow* F (or R), *emissions* E and the *share of the fossil sector* Γ. This graphic is published in [Kit2017e] and reused here with slight modifications.

oceans, plants and soil. To keep the complexity of the model as low as possible, we do not explicitly model a carbon cycle as in [And2013] but simply assume the carbon uptake leads to an exponential relaxation towards equilibrium on a characteristic time scale of $\tau^{\mathrm{A}} \approx 50\,\mathrm{a}$ [a = years]. Hence our first model equation is

$$\frac{dA}{dt} = E - A/\tau^{\mathrm{A}}, \tag{4.39}$$

where E will be derived below from economic assumptions.

The second variable is *economic output / production* Y [US$/a] representing the relation to wealth of a society, using the gross world product as its indicator as usual. We assume the economy to have a positive basic growth rate $\beta \approx 3\,\%$ [1/a] and additional climate impacts as in [Kel2011]. As a proxy for temperature we simply use A, effectively assuming an infinitely fast greenhouse effect. Hence this term is represented by $-\theta AY$ where $\theta \approx 8.57 \cdot 10^{-5}\,/(\mathrm{GtC\,a})$ is a temperature sensitivity parameter chosen such that the total growth rate $\beta - \theta A$ becomes negative when A exceeds the level corresponding to a global warming of $+2\,^{\circ}\mathrm{C}$. This gives

$$\frac{dY}{dt} = \beta Y - \theta AY. \tag{4.40}$$

The third dynamical variable is the *renewable energy knowledge stock* S that indicates how much knowledge is available for the production of renewable energy R [GJ/a = giga joule per year]. In accordance with Wright's law (e.g., [Nag2013]) of learning-by-doing, we basically identify S with the past cumulative production of renewables and thus measure it in units of [GJ]. To account for the human capital component, we additionally assume that knowledge depreciates on a characteristic time scale of $\tau^{\mathrm{S}} \approx 50\,\mathrm{a}$. Cumulation and depreciation then give

$$\frac{dS}{dt} = R - S/\tau^{\mathrm{S}}, \tag{4.41}$$

where R will be derived below.

Finally, to determine E and R, we use the following simplistic economic assumptions. The energy demand U [GJ/a] is proportional to the economic output

$$U = Y/\psi, \tag{4.42}$$

where $\psi \approx 147\,\mathrm{US\$/GJ}$ is an energy efficiency parameter. This demand is satisfied by a mix of fossil and renewable energy which are assumed to be perfect substitutes (and ignoring other energy sources such as agriculture and other bioenergy). Their respective shares are determined by a price equilibrium. We assume convex monomial cost functions and unit costs of renewable energy that show a power-law decay with growing S [Nag2013]. This implies that the fossil sector has a share given by the sigmoidal function

$$\Gamma = \frac{1}{1 + \left(\frac{S}{\sigma}\right)^{\rho}}, \tag{4.43}$$

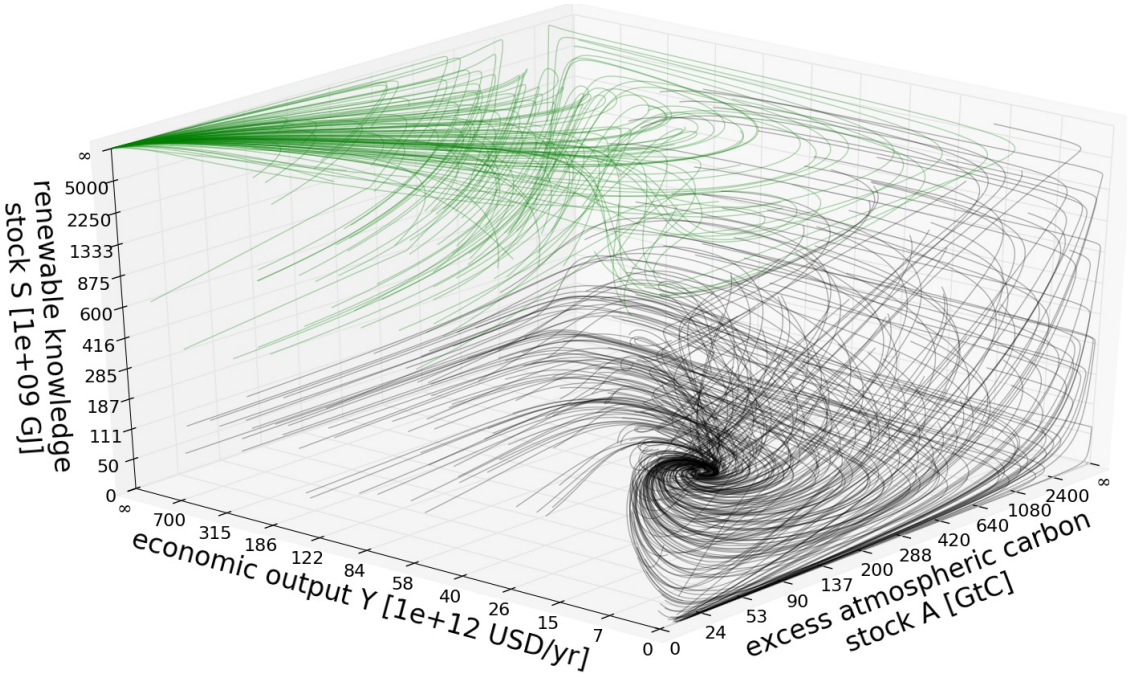


Figure 4.10: The default flow of the AYS model is sampled with trajectories from randomly distributed initial conditions on nonlinearly scaled axes so the full states space $\mathfrak{X} = \mathbb{R}^3_{\geq 0}$ is displayed (cf. Section 4.6.6). Green trajectories end up at the green attractor x^{g} and black ones at x^{b} which are analyzed in Section 4.6.2. This graphic was published in [Kit2017e].

where $\sigma \approx 4 \cdot 10^{12}\, GJ$ is the break-even knowledge level at which renewable and fossil extraction costs become equal, and $\rho \approx 2$ is a dimensionless parameter determined from the cost convexity and learning rate. Γ approaches unity (no renewables) as $S \to 0$ and zero (no fossils) as $S \to \infty$. Fossil and renewable energy flows and emissions are then

$$F = \Gamma\, U, \qquad R = (1 - \Gamma)\, U, \qquad E = F/\phi, \tag{4.44}$$

where our final parameter $\phi \approx 4.7 \cdot 10^{10}\,\mathrm{GJ/GtC}$ is the fossil fuel combustion efficiency.

This completes the model equations. Appendix B.2 contains details on how we estimated the parameters. The 3 dynamical variables A, Y, and S, combined to a vector $x = (A, Y, S)^\top$, are interrelated due to the various connecting equations and the nonlinearities arise particularly due to Equation (4.43) and Equation (4.40). The resulting flow is depicted in Figure 4.10 where the basins of the two attractors (discussed in Section 4.6.2) are already colored differently.

In the following, we will refer to these parameters as the *default parameters* and the corresponding flow as the *default flow*.

4.6.2 Attractors

With the above parameter values, the dynamics has two fixed points. The "black fixed point" x^{b} at

$$x^{\mathrm{b}} = \begin{pmatrix} A^{\mathrm{b}} \\ Y^{\mathrm{b}} \\ S^{\mathrm{b}} \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\theta} \\ \frac{\phi\psi\beta}{\theta\tau^{\mathrm{A}}} \\ 0 \end{pmatrix} = \begin{pmatrix} 350\,\mathrm{GtC} \\ 4.84 \cdot 10^{13}\,\mathrm{US\$} \\ 0 \end{pmatrix} \tag{4.45}$$

corresponding to a carbon based ecomony without renewable energy use, reduced economic output, and constant climate damages. And there is the "green fixed point" x^{g} at

$$x^{\mathrm{g}} = \begin{pmatrix} A^{\mathrm{g}} \\ Y^{\mathrm{g}} \\ S^{\mathrm{g}} \end{pmatrix} = \begin{pmatrix} 0 \\ +\infty \\ +\infty \end{pmatrix} \tag{4.46}$$

corresponding to eventually unbounded exponential growth of economic output and renewable knowledge. Also, there is an exponential decline of fossil usage and emissions towards zero. Both attractors can be identified in Figure 4.10. The mathematical meaning of "$+\infty$" is made clear in Section 4.6.6.

Both attractors are rather extreme cases. We find this acceptable because this model is a first example and we want to focus on the transients. We understand the asymptotics to be conceptual.

4.6.3 Current State

The current state $x^{\mathrm{c}} = (A^{\mathrm{c}}, Y^{\mathrm{c}}, S^{\mathrm{c}})$ can be estimated. A^{c} is currently around 240 GtC, corresponding to a concentration of 400 ppm [Bet2016], and the world gross product of 2015 is around 70 Trillion US$ [The2016b]. S^{c} is estimated on the basis of the total past renewable energy consumption of roughly $2 \cdot 10^{12}$ GJ [Mon2005]. Since this figure has accumulated over roughly the same time as the characteristic knowledge depreciation time, $\tau^{\mathrm{S}} = 50$ a, we assume roughly half of it has already depreciated, leaving 10^{12} GJ. Because of the large error margins involved in estimating this figure and because it contains hydroelectricity whose growth potential is somewhat problematic, we aim at staying on the conservative side with our estimate and again take only half of this value, giving $5 \cdot 10^{11}$ GJ.

$$x^{\mathrm{c}} = \begin{pmatrix} A^{\mathrm{c}} \\ Y^{\mathrm{c}} \\ S^{\mathrm{c}} \end{pmatrix} = \begin{pmatrix} 240\,\mathrm{GtC} \\ 7 \cdot 10^{13}\,\mathrm{US\$} \\ 5 \cdot 10^{11}\,\mathrm{GJ} \end{pmatrix} \tag{4.47}$$

4.6.4 Desirable States

The planetary boundary for climate change (PB-CC) is set to 350 ppm [ppm = parts per million] with an uncertainty zone until 450 ppm [Ste2015b]. We chose the desirable states

to be where $A > A^{\mathrm{PB}} = 345\,\mathrm{GtC}$ (above the pre-industrial level A^0), corresponding to the looser 450ppm boundary (see Appendix B.2 for the conversion).

Raworth demands that the planetary boundaries are to be complemented with social boundaries that she calls *social foundations* [Raw2012]. Combining both ideas is important to have not only a safe operating space but also a just operating space. Her *social foundations* shaped only illustrative indicators so we have to choose a suitable one for this case. A pragmatic choice is a lower boundary of $Y^{\mathrm{SF}} = 4 \cdot 10^{13}\,\mathrm{US\$}$ (SF-Y), the economic production of the year 2000. The exact value is open for discussion, but as our model has rather low complexity this choice seems reasonable. When trying to refine the number, one should model the distributions in order to include distributive justice and tackle inequality.

Note that for the default dynamics, the green fixed point $\boldsymbol{x}^{\mathrm{g}}$ is not violating either of the boundaries, while the black one $\boldsymbol{x}^{\mathrm{b}}$ violates the PB-CC as can be seen in Figure 4.11 (a).

4.6.5 Management Options

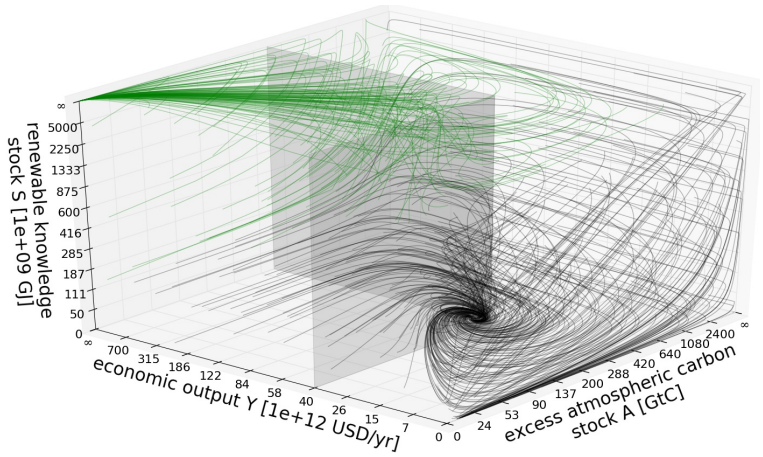
The above parameter values define what we consider the *default dynamics* since they represent a "business-as-usual" case. This means humanity applies no specific management that would alter "the way things usually go".

In addition to the default dynamics, we study some *management options* representing possible policy choices that may be combined in any way, leading to more or less shifted trajectories.

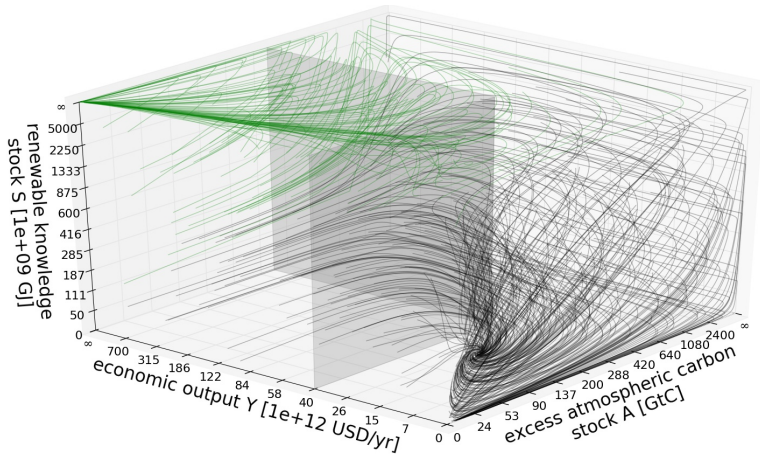
(i) The option of *low growth* (LG) reduces the basic growth rate β to half its value $\beta^{\mathrm{LG}} = 1.5\%/\mathrm{a}$. This moves the black fixed point to $\boldsymbol{x}^{\mathrm{b,LG}} = (A, Y, S) = (175\,\mathrm{GtC}, 2.42 \cdot 10^{13}\,\mathrm{US\$/a}, 0)$, no longer violating the PB-CC (see Figure 4.11 (b)) but now violating the SF-Y.

(ii) Climate mitigation by inducing an *energy transformation (ET)*, e.g. via taxing fossils and/or subsidizing renewable resource use. These policy instruments shift the relative costs of fossil and renewable energy, which according to Equation (4.43) can be effected in our model by a reduction of σ. Hence, we represent this option by reducing σ to approx. $(1/2)^{1/\rho} = 1/\sqrt{2}$ of its default value, i.e. to $\sigma^{\mathrm{ET}} = 2.83 \cdot 10^{12}\,\mathrm{GJ}$, corresponding to dividing the renewable to fossil cost ratio by half. This does not affect the location of the two attractors. But, more important, it changes the shape of the basins of attractions. When carefully inspecting Figure 4.11 (c), one can see that the volume of the green fixed point's the basin of attraction is enlarged in comparison to the default flow in Figure 4.11 (a). Within the concept of Basin Stability [Men2013; Men2014] the volume of the basin of attraction has been found to be an important indicator for an attractor's stability hence we will use a similar approach for the bifurcation analysis in Section 4.6.8.

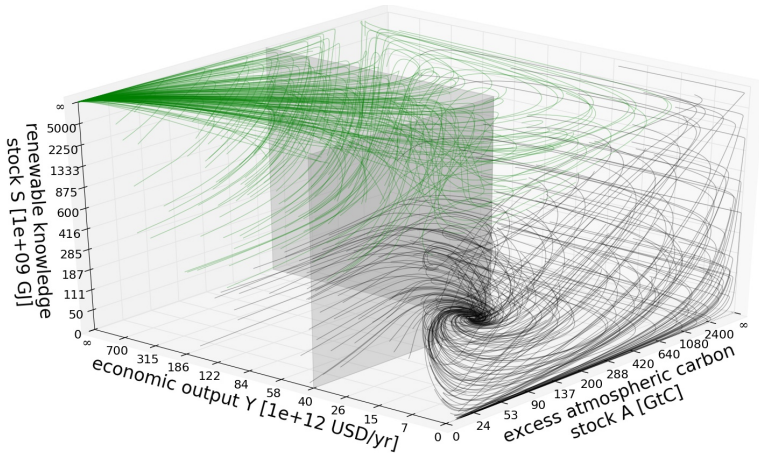
In terms of the control system as used for the variant definition of TSM in Section 4.3, we have the *default control* $\boldsymbol{u}^{\mathrm{d}} = (\beta, \sigma)$ and the two *management controls* $\boldsymbol{u}^{\mathrm{LG}} = (\beta^{\mathrm{LG}}, \sigma)$ and $\boldsymbol{u}^{\mathrm{ET}} = (\beta, \sigma)$. Together, they compose the set of controls $\mathcal{U} = \{\boldsymbol{u}^{\mathrm{d}}, \boldsymbol{u}^{\mathrm{LG}}, \boldsymbol{u}^{\mathrm{ET}}\}$.



(a) Default dynamics



(b) Low growth (LG)



(c) Energy transformation (ET)

Figure 4.11: The flows of the AYS model for (a) the default dynamics, (b) the low growth (LG) and (c) the energy transformation management mangement (ET) option including the combination of both boundary. Note how the black attractor x^{b} changes its position as discussed in Section 4.6.5. Again, the axes are nonlinearly scaled so the full states space $\mathcal{X} = \mathbb{R}^3_{\geq 0}$ is displayed (cf. Section 4.6.6). These graphics were published in [Kit2017e].

4.6.6 Dealing with the Unbounded State Space

The model (4.39)-(4.44) in Section 4.6.1 is defined on an unbounded state space and in this section we map it to a bounded one for two reasons: the attractor at "$\boldsymbol{x}^{\mathrm{g}} = (0, +\infty, +\infty)$" and the need for a bounded state space in order to apply the Saint-Pierre algorithm (SPA). In order to make this mathematically sound, we map $\mathcal{X} = \mathbb{R}^3_{\geq 0}$ (parameterized by the variables $\boldsymbol{x} = (A, Y, S)$) to a bounded space $\mathcal{Y} = [0, 1)^3$ (parameterized by transformed coordinates $\boldsymbol{y} = (a, y, s)$) and then add the point $\boldsymbol{y}^{\mathrm{g}} = (0, 1, 1)$ which is the equivalent of $\boldsymbol{x}^{\mathrm{g}}$ in these new coordinates. So we perform a change of coordinates

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}) \quad \longrightarrow \quad \dot{\boldsymbol{y}} = \tilde{h}(\boldsymbol{y}), \tag{4.48}$$

where we switch from the old right-hand side (RHS) f to the new RHS F. Following the explanations in Section 4.5.3, we use the transformation

$$\begin{aligned} \Phi : \mathcal{X} = [0, \infty)^3 &\longrightarrow \mathcal{Y} = [0, 1)^3 \\ A &\longmapsto a = \tfrac{A}{A^{\mathrm{mid}} + A} \\ Y &\longmapsto y = \tfrac{Y}{Y^{\mathrm{mid}} + Y} \\ S &\longmapsto s = \tfrac{S}{S^{\mathrm{mid}} + S}, \end{aligned} \tag{4.49}$$

where the parameters $\boldsymbol{x}^{\mathrm{mid}} = (A^{\mathrm{mid}}, Y^{\mathrm{mid}}, S^{\mathrm{mid}})^\top$ are such that $\Phi(\boldsymbol{x}^{\mathrm{mid}}) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^\top$. They can be understood as the scale where the "resolution is the best". Hence, changing this value does not qualitatively influence the result, but a good choice can make them clearer.

This is exactly the transformation that has been used to create Figures 4.10 and 4.11 (a) to 4.11 (c) and will be used for all the following figures, too. As we care most about the current state of the world, we choose $\boldsymbol{x}^{\mathrm{mid}} = \boldsymbol{x}^{\mathrm{c}}$.

Using Equation (4.34), we get a new set of ODEs with $\boldsymbol{y} = (a, y, s)$ as coordinates

$$\dot{a} = \frac{Y^{\mathrm{mid}}}{\phi\psi A^{\mathrm{mid}}}\gamma(1-a)^2\frac{y}{1-y} - \frac{a(1-a)}{\tau^{\mathrm{A}}}, \tag{4.50a}$$

$$\dot{y} = y(1-y)(\beta - \theta A^{\mathrm{mid}}\frac{a}{1-a}), \tag{4.50b}$$

$$\dot{s} = (1-\gamma)\frac{Y^{\mathrm{mid}}}{\psi S^{\mathrm{mid}}}(1-s)^2\frac{y}{1-y} - \frac{s(1-s)}{\tau^{\mathrm{S}}}, \tag{4.50c}$$

$$\gamma = \frac{(1-s)^\rho}{(1-s)^\rho + \left(\frac{S^{\mathrm{mid}} s}{\sigma}\right)^\rho}, \tag{4.50d}$$

where γ is the equivalent of Γ in Equation (4.43) but in the $\boldsymbol{y}$-coordinates. The fixed

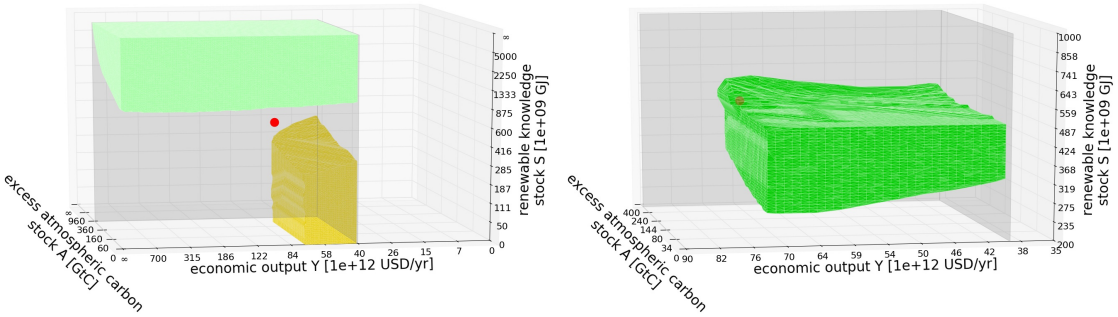


Figure 4.12: The safest region, the Shelter (right, light green), is the part where enough knowledge on renewable resource use has already been accumulated and hence the relative price of the renewable resource is low enough such that even the exponentially growing total energy demand which is proportional to the economic output can be fully compensated by the renewable energy resource so the planetary boundaries are not transgressed. The Backwater (left, yellow) corresponds to the region where one can stay using the *low growth* management option leading finally to a *zero-growth economy* but still staying within the boundaries. These graphics were published in [Kit2017e].

points in the new y-coordinates are

$$\boldsymbol{y}^{\mathrm{g}} = \begin{pmatrix} a^{\mathrm{g}} \\ y^{\mathrm{g}} \\ s^{\mathrm{g}} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \qquad \longleftrightarrow \qquad \boldsymbol{x}^{\mathrm{g}} = \begin{pmatrix} A^{\mathrm{g}} \\ Y^{\mathrm{g}} \\ S^{\mathrm{g}} \end{pmatrix} = \begin{pmatrix} 0 \\ \infty \\ \infty \end{pmatrix} \tag{4.51a}$$

$$\boldsymbol{y}^{\mathrm{b}} = \begin{pmatrix} a^{\mathrm{b}} \\ y^{\mathrm{b}} \\ s^{\mathrm{b}} \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\beta+\theta A^{\mathrm{mid}}} \\ \frac{\phi\psi\beta}{\phi\psi\beta+Y^{\mathrm{mid}}\theta\tau^{\mathrm{A}}} \\ 0 \end{pmatrix} \qquad \longleftrightarrow \qquad \boldsymbol{x}^{\mathrm{b}} = \begin{pmatrix} A^{\mathrm{b}} \\ Y^{\mathrm{b}} \\ S^{\mathrm{b}} \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\theta} \\ \frac{\phi\psi\beta}{\theta\tau^{\mathrm{A}}} \\ 0 \end{pmatrix}. \tag{4.51b}$$

Now, we formally extend the dynamics such that $\tilde{h}(\boldsymbol{y}^{\mathrm{g}}) = 0$. As we can observe from the flow plots in Figures 4.11 (a) to 4.11 (c) that $\boldsymbol{y}^{\mathrm{g}}$ is a a fixed-point-like attractor, this is just a formal definition to include that point in the dynamics. Note that the rest of the boundaries, where $y = 1$ or $a = 1$, are not easily included in the dynamics, because the flow diverges on these boundaries.

4.6.7 Model Analysis Results

The model introduced in Section 4.6.1 has been analyzed in its compactified form Equations (4.50a) to (4.50d). Then, we used the nonlinear, local time-homogenization from Section 4.5.4 and the Saint-Pierre algorithm (SPA) that was sketched in Section 4.5. We do not write out the equations for the time-homogenized version as they are lengthy and their calculations.

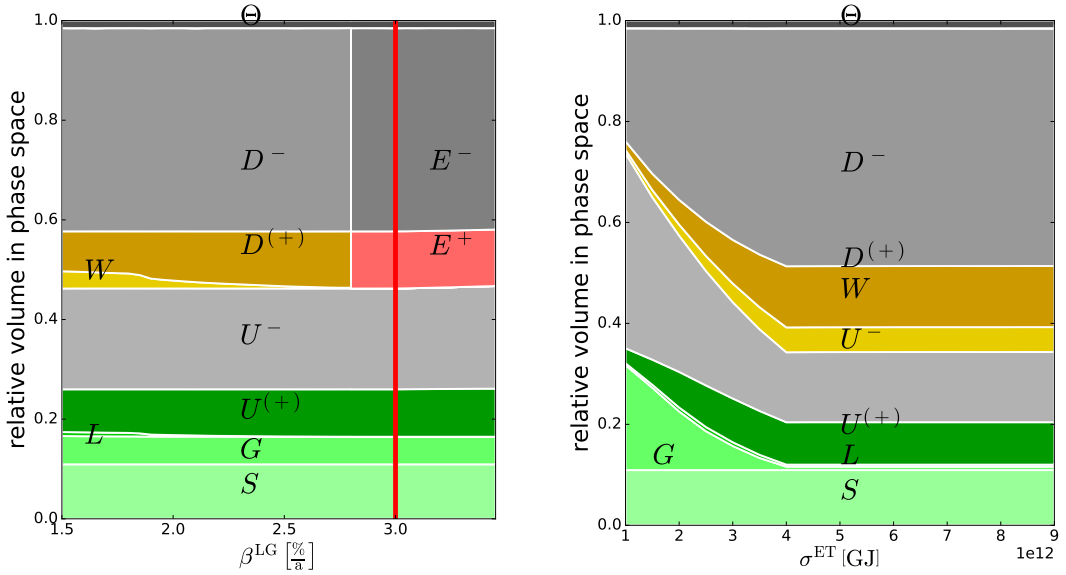
The most important identified regions are depicted in Figures 4.12 (a) and 4.12 (b) and we use them for the following discussion.

The first regions to note are the *shelters* and *backwaters*, depicted in Figure 4.12 (a). The former, where one can stay without management forever in the sun and which is the

safest region thus, is in our model the invariant kernel of the green fixed point's basin of attraction when restricting to the desirable states $\mathfrak{X}^+$ only. When being in $\mathfrak{X}^+$ and having accumulated already enough knowledge for the energy production with renewable resources, they become so cheap that there is basically no need for fossil fuels anymore. So the remaining (excess) CO_2 (above long-term equilibrium) is removed over time due to the carbon uptake, leading the system to the green fixed point. The *glades*, where one can reach the shelter through the desirable region, are just a thin layer under the shelter so they have not been included in Figure 4.12 (a).

The *backwater*, where one can stay in the sun forever but needs to apply management over and over again, is the part of the desirable region where the growth of economic output and hence of emissions can be restricted. That way, the atmospheric carbon concentration can be kept within the planetary boundary. Also, within the backwater the decarbonization of the economy is impossible, since the given maximal carbon tax and renewable subsidy policy are too weak to make renewables competitive with fossil fuel. Instead, one can manage to stay in a state that corresponds to a carbon-based economy where the low background economic growth is compensated by the climate impacts. The atmospheric carbon level is relatively high but still within the boundary and in equilibrium with the emissions. Hence, using the low growth option properly, one can stay within the desirable region but cannot reach the green fixed point. Note that for simplicity, we did not include the option to choose a value of the base growth rate lying between the two options β and β^{LG}. So formally, the management strategy required to stay in the desirable state described above involves a fast switching between β and β^{LG}, since either of these two extreme values alone leads to a black fixed point in the undesirable region. Still, it is easy to see that this management strategy is equivalent to using a constant intermediate value of β^{m} (e.g. $\beta^{\mathrm{m}} = 2.7\,\%/\mathrm{a}$) instead. This dynamics has a black fixed point that lies in the desirable region. So this real-world option is implicitly included in the model. In other words, since the TSM-framework allows arbitrarily fast and frequent switches between the management options, one only needs to model the "corners" of the option space explicitly and gets all intermediate options (= all convex combinations). Still, replacing the value of β^{LG} with β^{m} would introduce further changes, since the maximal management in the transient would be restricted and thus the size of the backwaters reduced. We show this in Section 4.6.8. Here, it becomes obvious that the TSM-framework incorporates the asymptotics and the transient of a model.

The current state $\boldsymbol{x}^{\mathrm{c}}$ (as estimated in Section 4.6.3) seems to lie between the two regions discussed above. It is in a region we call a *time-limited lake* as can be seen in the zoom in Figure 4.12 (b). This means humanity appears to be currently facing the so-called lake dilemma where we have to make the qualitative decision between staying within the desirable region uninterruptedly but being in need of management forever or going through the undesirable region to finally end up in the shelter. In the model, the choice is between using the energy-transformation option (ET) in order to speed up the knowledge accumulation on renewable resource use and finally reach the green fixed point or the low growth option (LG) in order to restrict the total energy demand and reach the black fixed point. Note that even with a combined usage of the different options (convex



(a) Low growth

(b) Energy transformation

Figure 4.13: The TSM-bifurcation diagrams when varying the parameters governing the two management options from their maximally deviating value to their corresponding value of the default dynamics (except (a) where the default dynamics value is marked by a red vertical line). In (a) a *Downstream-Eddies-Bifurcation* occurs when varying β^{LG} due to the stable focus of the low growth option crossing the boundary between desirable and undesirable states. In (b), we observe that for small σ^{ET}, i.e. strong management, the glade and dark upstream increase in size drastically because the meaning of the energy transformation option is to make the transition to the renewable resource use easier. These graphics were published in [Kit2017e].

combinations) it cannot be avoided to transgress the boundary when going for the first choice.

4.6.8 Bifurcation Analysis

Within the example model from Section 4.6.1, varying the parameters corresponding to the two managements options may lead to bifurcations because of possible changes in the topological structure of the state space with respect to TSM. We call these *TSM-bifurcations*. As an indicator for the TSM-bifurcations we use the relative volume of each region, motivated by the concept of Basin Stability [Men2013; Men2014] and its extensions [Hel2016; Kan2016; Kit2017f; Kli2015; Mit2015; Mit2017a; Mit2017b]. Because we use uniformly distributed points in state space for the SPA, we estimate the relative volume of one region with the number of points associated to this region over the total number.

When varying β^{LG} in Figure 4.13 (a) corresponding to the *low growth* management option (LG) from 1.5 %/a to 3.5 %/a, a *downstream-eddies bifurcation* occurs. Until the

fixed point of the LG flow crosses the planetary boundary at the critical value of $\beta^{\mathrm{LG}} = \beta^{\mathrm{PB}} \approx 2.95\%/\mathrm{a}$ there is always a *backwater*. Beyond, there are only *eddies* left. The eddies occur because the focus of the default flow and the one of the log-growth flow are both in the undesirable region. One can (in this case) switch between the two flows in a smart way such that one circles far around both foci and can reach the desirable region over and over again while having to pass through the undesirable region in between. Due to the discretization in state space and time during the estimation, the bifurcation seems to occur already at $\approx 2.8\%/\mathrm{a}$ in Figure 4.13 (a).

As discussed in Section 4.6.7 the *backwater* for $\beta^{\mathrm{LG}} = 1.5\,\%/\mathrm{a}$ occurs because there exists an in-between value $\beta^{\mathrm{m}} = 2.7\,\%/\mathrm{a}$ where the focus lies in the desirable region. In Figure 4.13 (a), the volume of the backwater for $\beta^{\mathrm{LG}'} = \beta^{\mathrm{m}} = 2.7\,\%/\mathrm{a}$ is smaller than for $\beta^{\mathrm{LG}} = 1.5\,\%/\mathrm{a}$, because the maximal management in the transient is restricted.

In Figure 4.13 (b), the change in σ^{ET} from the *energy-transformation* management option is depicted and a notable change for small σ^{ET} corresponding to strong management can be observed. As the name implies, this management speeds up the transformation to renewable resource use and thus the glade increases in size, because from more initial conditions it is possible to reach the green fixed point without transgressing the boundaries.

4.7 Summary

In this chapter, we have introduced the framework topology of sustainable management (TSM) to study models with respect to notions of desirable states and possible management influences. Furthermore, we have connected it to Viability Theory (VT) enabling us to automatize the analysis. We call this process *operationalization*.

Additionally, we have extended the framework with the distinction of lakes in finite-time and infinite-time lakes. At the same time, this has naturally led to the pressing lake dilemma, an aggravated version of the lake dilemma where the corresponding decision has to be made in limited time.

We have analyzed multiple 2-dimensional examples in order to illustrate the vast range of topics where the framework is applicable. They have ranged from natural, co-evolutionary, socio-economic to mechanical. In the global carbon cycle, we have observed that even though the final fixed-point of the default flow was always reachable, a rich topology emerged due to the need for management at certain points. In the combined population and resource dynamics, we have seen that simple parameter changes could induce vast changes of the topology, even thought the flows and borders structurally stayed the same. In the gravity pendulum, our example from classical mechanics, a wider variety could be observed, including a natural occurrence of eddies and a lake, the latter leading to the lake dilemma. The competing plant types have lead to lake dilemma that is like a weak version of a lock-in effect. If one uses plant type 1 only, one has to manage forever to ensure staying in the desirable region. But by changing the management and planting type 2, the farmer can change the soil properties. Still, for some time he cannot

be in the desirable region, but finally his focus on plant type 2 will let reach reach the shelter. In the "substitution of a dirty technology" model, a proper lock-in effect has been observed. If one focuses too much on dirty technology, there is a point where one cannot leave the undesirable region anymore. Hence, an abyss and a trench have occurred.

Taking the step from 2 dimensions to 3 induces the need to automatically analyze a model. For that, we have introduced TSM based on VT and presented the Saint-Pierre algorithm (SPA). Because it works on bounded spaces only, we have introduced a coordinate transformation to a bounded space. The transformation parameters are chosen to fix the scale of "highest resolution". Furthermore, we have solved the problem that time scales may vary through orders of magnitude by introducing a nonlinear, local time homogenization. Additionally, we have introduced an extension of the SPA for implicit capture basins.

These concepts have been applied to a 3-dimensional example system combining climate change, economic output and energy transformation. The system has been kept minimalistic, to enable our analysis, so while the results are interesting, they should be understood with caution. A rich topology was found and the current state of the world has been estimated within the scope of the model; particularly interesting was that it seems to be inside a finite-time lake. So, humanity might be facing a pressing lake-dilemma where it has to make a choice between two qualitatively different options. Furthermore, we have performed a TSM-bifurcation analysis under change of the management parameters and found a downstream-eddies bifurcation.

Chapter 5

Global Teleconnectivity Structures of the El Niño–Southern Oscillation and Large Volcanic Eruptions

In Chapters 3 and 4, methods for the transient analysis have been developed and applied to multiple example models. A natural next step is to adapt these methods to analysis of time series. A possible approach is to understand a strong perturbation from the "normal state" within a time series as "setting the initial condition" of a trajectory. That way, the return to the normal state could be interpreted as the transient phase of the trajectory. This idea is based on the motivation of Basin Stability in [Men2013]. Unfortunately, this approach did not work out. Still, the following analysis of strong perturbations within the climate system, in particular El Niño and La Niña events as well as volcanic eruptions, have lead to this chapter in my thesis.

The analysis of climate data is fundamental to understand climate phenomena, e.g. El Niño, and develop more accurate methods for forecasts. Typically, these data sets are spatially resolved, i.e. gridded, time series of observables such as temperature, precipitation, pressure or other climate variables. While the analysis of such big climate data sets has been traditionally attempted by means of classical statistical approaches like empirical orthogonal function or maximum covariance analysis [Sto2003], these methods exhibit fundamental intrinsic limitations, including their linearity and associated condition of pairwise orthogonal patterns [Gám2004].

During the last years, complex network representations of climate variability have been developed [Don2009a; Don2009b; Ste2012; Tso2006; Tso2011] and demonstrated to provide a suitable approach for relieving some of the aforementioned concerns [Don2015]. This approach is referred to as *functional climate network analysis*. The individual grid points are considered as nodes of a spatially embedded networks; links among these nodes are established according to similarities between the individual (local) climate time series [Don2009a; Don2017; Tso2006]. Therefore, the obtained network structures highlight essential statistical interrelationships among spatio-temporal climate data [Don2009a].

One rather fundamental property of large networks is their (possibly hierarchical)

organization in terms of communities – an aspect that was also addressed recently in the context of climate data analysis [Tso2011]. A community is a subset of densely connected nodes which exhibit only few interactions with the rest of the network [For2010; New2006a]. In this context, Tsonis et al. [Tso2011] argued that each community in a climate network can be considered as a subsystem which operates relatively independent of the other communities. Besides corresponding connectivity structures in individual climate variables, community detection algorithms [For2010] can also be used to detect multi-variable clusters [Ste2010].

In this chapter, we analyze global surface air temperature (SAT) data in terms of functional climate networks. We demonstrate the close relationship between El Niño and La Niña episodes as well as strong volcanic eruptions with temporal changes in the modular organization of the resulting networks. For this purpose, we study the teleconnectivity structure in the climate system in terms of spatial fields of two network characteristics: the degree, that represent the number of strong statistical connections, and the average distance of a link. In addition, the associated temporal variations are traced by scalar-valued global network characteristics.

The chapter is structured as follows: After this introduction follows a short section on Code availability. Then, Section 5.2 gives a brief climatological background of ENSO and volcanic eruptions. The data and methodology used in this work are described in detail in Section 5.3. Finally, the results are presented and discussed in Section 5.4.

This chapter is based on the publication "Global teleconnectivity structures of the El Niño–Southern Oscillation and large volcanic eruptions: An evolving network perspective" [Kit2017d].

5.1 Code Availability

All code used in this work has been written in Python and published under GPLv3 license as a GitHub repository [Kit2017a]*. Detailed information for the reproduction of the results in this chapter can be found there. The published code was originally designed to produce these specific result, but we kept it rather general with further extensions in mind. Thereby, it can be a starting point for future evolving network research. It provides some basic structures that are needed for evolving network analysis, an interface for `hdf5` (a high-performant file type for data storage) and automatic parallelization using `mpi`. In case of questions, please contact the author of the thesis.†

5.2 Climatological Background

5.2.1 El Niño–Southern Oscillation

The El Niño–Southern Oscillation (ENSO) is on of the dominant phenomena within the climate system due to the magnitude of associated variations in sea-surface temperatures

*`https://github.com/timkittel/global-structures-enso-volcanoes`

†`Tim.Kittel@pik-potsdam.de`

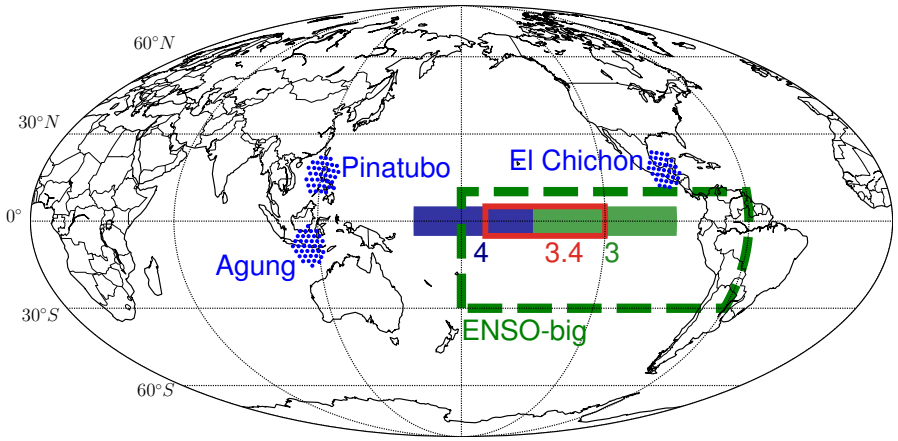


Figure 5.1: Main regions of interest used within this chapter. Sets of blue dots labeled with "El Chichon", "Agung" and "Pinatubo" indicate grid points within a 5° radius around the corresponding volcanoes. The numbers 3, 3.4 and 4 identify the corresponding Nino regions (cf. Table 5.1) commonly used for defining characteristic indices of ENSO variability. The region "ENSO-big" will be removed from the complete global data set when analyzing the spatial imprints of volcanic eruptions to ensure that ENSO-related effects are excluded. This graphic was published in [Kit2017d].

region	latitudes	longitudes
Nino1+2	10°S - 0°N	90°W - 80°W
Nino3	5°S - 5°N	150°W - 90°W
Nino4	5°S - 5°N	160°E - 150°W
Nino3.4	5°S - 5°N	170°W - 120°W
ENSO-big	30°S - 10°N	180°W - 60°W

Table 5.1: Overview on different regions commonly used for defining characteristics temperature-based indices associated with ENSO variability. In addition, we include the definition of the "ENSO-big" region studied in this work, which corresponds to the region that is discarded in our analyses of the impacts of strong volcanic eruptions on global temperature teleconnectivity. This table was published in [Kit2017d].

and sea-level pressure [Tre1997]. During the positive phase (El Niño) of this complex oscillation of the coupled atmosphere-ocean system in the Tropical Pacific, the Eastern Tropical Pacific exhibits some anomalous warming with respect to "normal" mean conditions. The negative phase (La Niña) is characterized by a corresponding cooling. It has been shown that effects of both ENSO phases can be observed in remote regions including North and South America, Africa, the Indian subcontinent, and even Antarctica [Cla2008; Dai2000; Nee2003; Rop1987; Sar2010; Tur2004].

The long-term variability of ENSO is characterized by some irregular oscillations with a period between 2 and 7 years. Following the prominent anomaly structures of Tropical sea surface temperature (SST) and sea level pressure fields, ENSO is commonly traced by indices that take up the variability of the aforementioned observables in some key region of the Tropical Pacific ocean. For example, a set of indices has been defined in terms of average SST anomalies taken over distinct regions in the Eastern and Central Tropical Pacific, referred to as Nino1+2, Nino3, Nino4 and Nino3.4 respectively [Tre2001] (see Figure 5.1 and Table 5.1). We will use the so-called Oceanic Niño Index (ONI) for differentiating between different phases of ENSO. The ONI is defined as the running three-month mean SST anomaly for the Niño 3.4 region. When the ONI exceeds 0.5°C for at least five consecutive months, the corresponding situation is classified as an El Niño. If the ONI drops below −0.5°C for at least five consecutive months, it indicates a La Niña episode. The strength of such an event is classified as weak, moderate, strong, or very strong if the absolute value of ONI exceeds 0.5°, 1.0°, 1.5°, or 2.0° respectively.

Particularly during the El Niño phases, but also during La Niña, a finer distinction in two types is usually been made as the SST anomaly in the Pacific region exhibits different patterns [Wie2016]. The first type is the classic or Eastern Pacific (EP) El Nino [Har1998; Ras1982]. As the name indicates, it is localized in the Eastern part of the Pacific and characterized by strong positive SST anomalies in front of the Western coast of South America. The second type is the El Niño Modoki or Central Pacific (CP) El Niño with the strongest signatures of SST anomalies in the central Pacific around the dateline. Similarly, both types, EP and CP, exist for La Niña, too. As there are some contradictory classifications in the literature, Wiedermann et al. [Wie2016] reviewed recent studies and presented a new indicator based on functional climate networks that we follow here (see Table 5.2).

5.2.2 Volcanic eruptions

Besides distinct ENSO episodes and their known global climate impacts, another type of events, that can substantially affect climate at large spatial and temporal scales, are strong volcanic eruptions. Similar to El Niño and La Niña episodes, such events can result in large-scale spatially coherent cooling trends. This happens due to modifications of the radiation balance by changes in atmospheric chemistry and the shielding effect of volcanic aerosols in the stratosphere. Such cooling can again cause large-scale changes of precipitation and temperature patterns from synoptic (weather) time scales up to even relatively persistent multi-annual effects [Rob2000].

Event	Years
EP El Niño	1957, 1965, 1972, 1976, 1982, and 1997
CP El Niño	1953, 1958, 1963, 1968, 1969, 1977, 1979, 1986, 1987, 1991, 1994, 2002, 2004, 2006, and 2009
EP La Niña	1964, 1970, 1973, 1988, 1998, 2007, and 2010
CP La Niña	1954, 1955, 1967, 1971, 1974, 1975, 1984, 1995, 2000, 2001, and 2011

Table 5.2: Summary of the classifications of ENSO events in the four flavors Eastern Pacific (EP) El Niño, Central Pacific (CP) El Niño, EP La Niña, and CP La Niña following [Wie2016]. This table was published in [Kit2017d].

Here, we focus on the global effects of the three major volcanic eruptions during the second half of the 20th century. The largest and most influential eruption, the Mount Pinatubo eruption [McC1995], took place between April and September 1991 in the Philippines, followed by the Mount Agung eruption in Indonesia (February 1963 to January 1964) and the El Chichon eruption (March to September 1982) in Mexico (for locations see Figure 5.1).

5.3 Data and Methods

5.3.1 Data

We use daily mean surface air temperature (SAT) data (at sigma level $\sigma = 0.995$) from the National Center for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR) Reanalysis I project [Kal1996; Kis2001]. The temporal range is from year 1948 to 2015. It is spatially resolved using a global grid with equi-angular resolution of 2.5° in both latitudes and longitudes, hence being comprised of $10,512$ individual temperature time series. Note that we found the inclusion of the land areas important. Hence we chose SAT instead of the aforementioned sea surface temperature (SST). In order to remove leading order effects of seasonality in the temperature records, the long-term average temperatures for each calendar day of the year are subtracted from the raw data independently for each grid point.

Equi-angular gridded data have by construction a higher density of grid points at the poles than around the equator. If not properly accounted for, this could result in systematic biases of statistical characteristics overemphasizing the polar regions with apparently more data. For the latter purpose, area-weighted measures have been developed and subsequently applied in recent studies [Hei2012; Tso2006; Wie2013]. As an alternative, we follow here the approach of Radebach et al. [Rad2013], where the original data is remapped onto a grid with a much higher spatial homogeneity, specifically the icosahedral grid as described by Heikes et al. [Hei1994]. This leads to a decomposition of the Earth's

surface into Voronoi cells of almost the same area, with $N = 10,242$ grid points that exhibit a narrowly peaked distribution of geodesic distances between direct neighbors.

In [Rad2013], the time series associated with each new grid point have been determined by the average of the respective four surrounding grid points in the original equi-angular grid. However, we use the four closest points in space instead, since the latter are likely to provide a more closer approximation of the climate variability at the new grid point. Furthermore, they can be determined efficiently using spatial search trees. Due to the commonly rather high spatial correlation length of the SAT field and its resulting spatial smoothness, we do not expect the time series resulting from both algorithmic variants to differ markedly.

Finally, we note that when using the global data set as described above, the temporal correlations associated with the key ENSO region and the surrounding parts of the Pacific ocean are known to dominate climate variability globally. This leads to undesired outcomes when aiming to resolve the effects of individual volcanic eruptions on global temperature patterns, since they might be masked by ENSO variability. In order to account for this problem, we are going to use the full set of data when studying the effects of ENSO on global temperature teleconnectivity, while excluding the main ENSO region and its surroundings (referred to as "ENSO-big" in Figure 5.1 and Table 5.1) when studying the impacts of volcanic eruptions. Note that this excluded region has been chosen rather large on purpose to ensure an as complete as possible separation between the direct ENSO impacts and the effects of the volcanic eruptions, especially in cases of simultaneous events.

5.3.2 Functional Climate Network Analysis

Functional climate networks are a coarse-grained spatial representation of the co-variability structure of some climate variable, SAT in this case [Don2017; Tso2006]. The geographical positions associated with each time series is identified as the nodes. Links are established if the pairs of time series of two nodes fulfill a certain similarity criterion (see below). Hence, these links represent the statistical association between climate variability at different points on the Earth's surface.

As introduced in Section 2.1, we conveniently represent networks in terms of their adjacency matrix A. $A_{ij} = 1$ indicates the existence of a link between node i and node j, while $A_{ij} = 0$ corresponds to its absence. In our specific case, the matrix A is time-dependent, since the spatial co-variability structure of the SAT field changes with time. In such a case, we speak of an *evolving climate network* [Rad2013].

Network Generation. We take the grid points of the icosahedral grid, constructed by remapping the original NCEP/NCAR reanalysis data, as nodes of an evolving SAT network, i.e. we consider a fixed node set that does not change over time.

For establishing the time-dependent link set, we consider sliding windows in time, covering a set of days $\{d\} = [d_0, d_0 + \Delta d]$ with a width of $\Delta d = 365$ days and mutual offset of 183 days between subsequent windows. Each of these windows is associated with

the corresponding midpoint $d_{mid} = d_0 + \Delta d/2$. From the seasonally adjusted temperature data at each grid point during a respective time window $T_i(\{d\})$, with i indicating the respective grid point or node, we compute the matrix of pairwise Pearson's correlation coefficients

$$c_{ij}(d_{mid}) = \frac{\langle T_i(\{d\})T_j(\{d\})\rangle_{\{d\}}}{\sqrt{\sigma_{\{d\}}(T_i(\{d\})) \cdot \sigma_{\{d\}}(T_j(\{d\}))}}\,, \tag{5.1}$$

where $\langle\bullet\rangle_{\{d\}}$ and $\sigma_{\{d\}}(\bullet)$ denote the mean value and standard deviation of the respective variable taken over the time window $\{d\}$. We identify the entries of c with the highest absolute values of mutual correlations, $|c_{ij}|$. Specifically, in this work, we consider the 0.5% strongest pairwise statistical similarities among all nodes per window, i.e.

$$A_{ij}(d_{mid}) = \Theta(|c_{ij}|(d_{mid}) - q_{|c|,0.995}(d_{mid}))\,, \tag{5.2}$$

where $\Theta(\bullet)$ is the Heaviside function and $q_{|c|,0.995}(d_{mid})$ is the 99.5-percentile of the distribution of absolute correlation values for the time window centered at d_{mid}. $A\,(d_{mid})$ is now the representation of the evolving climate network. By construction, A is symmetric corresponding to an undirected network.

Transitivity. As introduced in Section 2.1, the network transitivity quantifies how clustered the connectivity of a network is. In equations, it is expressed as (see Equation (2.3), here just repeated)

$$\mathcal{T} = \frac{\sum_{ijk} A_{ij}A_{jk}A_{ki}}{\sum_{ijk} A_{ij}A_{jk}}. \tag{5.3}$$

Wiedermann et al. [Wie2016] used this in the context of functional climate networks to create an indicator for discriminating EP and CP El Niño and La Niña events.

Modularity. The definition of the modularity $\mathcal{Q}$ is given in Section 2.1. It is a global measure of heterogeneity within the network. It describes how well different groups of nodes can be identified, that are densely connected within each group but only sparsely to other groups. In the case of a climate network, modularity provides a single scalar-valued characteristic property that discriminates between a relatively homogeneous link placement (low modularity) and the existence of (commonly regional confined) clusters of nodes (time series) that exhibit relatively coherent variability (high modularity).

While the definition of modularity in Equation (2.4) is mathematically precise, to find the maximum is a computationally hard problem and can only be estimated by suitable algorithms. Hence, several estimation algorithms have been proposed [For2010]. It should be emphasized that many of them can result in suboptimal solutions. Thus, a good choice of the algorithm is important for obtaining reliable results. We employ the *WalkTrap* method introduced by Pons et al. [Pon2006]. By comparing the results provided by this algorithm with those of other methods (see Appendix C), we find that the WalkTrap solution exhibits comparably high values of modularity and little background noise. This

indicates a relatively high degree of stability of the identified network communities in case of strongly overlapping windows, i.e. in cases where the data does not change much. Note that this result concerns the specific functional climate networks that we employ here. For other kinds of networks, a different choice of algorithm might be optimal.

Degree. As already introduced in Section 2.1, the degree k_i of a node i is defined as the number of links connected to i (see Equation (2.1), here just repeated)

$$k_i = \sum_{j=1}^{N} A_{ij} \,. \tag{5.4}$$

In case of a functional climate network, the degree can be considered as a proxy for the importance (or centrality) of a certain grid point in the spatio-temporal correlation structure of the variable of interest.

In the following, we refer to network measures like the degree, which provide a characteristic value specific to each individual node i, as *local network characteristics*. We call the full set of their values taken together with the associated spatial positions of all nodes a *field*. Hence, we just introduced the *degree field*.

Average Link Distance. A local network characteristic of a functional climate network is the average link distance of a node i [Don2017] which is defined as

$$l_i = \langle l_{ij} \rangle_{\{j|A_{ij}=1\}} = \frac{1}{k_i} \sum_j A_{ij} l_{ij} \,, \tag{5.5}$$

with the normalized spatial distance l_{ij} between two nodes i and j. The corresponding field is called *average link distance field*. As normalization, we choose the largest possible (shortest) distance between two points on the Earth's surface, i.e. half of the circumference of the Earth, u_{Earth}. Hence, $l_{ij} = 2L_{ij}/u_{Earth}$ where L_{ij} is the geodesic distance between nodes i and j. A low average link distance indicates that i has very localized connections, while a high value points to a node with global connections.

We emphasize that the average link distance is not to be confused with the average path length. The latter is the minimum number of links separating node i from j and then averaged over all possible nodes j.

Taking the average of l_i over all nodes i of the network gives the *global average link distance*

$$\langle\langle l \rangle\rangle = \langle l_i \rangle_i \,. \tag{5.6}$$

5.3.3 Regionalization of Field Measures

As detailed above, node degree and average link distance constitute two important local network characteristics. In some of our following investigations, it will be useful to study the associated spatial fields in full detail. However, when focusing on the specific impacts

of certain climate phenomena, it can be beneficial to perform a regionalization of these measures. Specifically, for a subset of nodes $\mathcal{L} \subseteq \mathcal{N}$, representing a certain part of the globe, a regionalized version of the degree is given by

$$k_{\mathcal{L}} = \frac{1}{|\mathcal{L}|} \sum_{i \in \mathcal{L}} k_i \,, \tag{5.7}$$

where $|\mathcal{L}|$ is the number of nodes in the set. As a consequence, we can not only assign a degree value to an individual node, but also (as a mean degree) to a subgraph. Note that this regionalized degree differs from the concepts of cross-degree and cross-link density between subgraphs [Don2011]. Unlike $k_{\mathcal{L}}$, the these exclude contributions due to links between nodes within $\mathcal{L}$ in their definitions.

For the average link distance, the corresponding regionalized property $l_{\mathcal{L}}$ is defined analogously.

Below, we detail some reasonable choices for $\mathcal{L}$ to be utilized in the context of the present work. They focus on specific spatially contiguous regions of the Earth's surface that are associated with ENSO or volcanoes with strong past eruptions.

5.3.3.1 El Niño–Southern Oscillation Regions

To measure the strength of an ENSO event, different indices such as the Ocean Niño Index (ONI) have been introduced. Therefore, 4 Niño regions (1+2, 3, 4 and 3.4) are defined [Tre2001]. We summarize their spatial extensions in Table 5.1 and depict them in Figure 5.1.

The regionalization procedure can be employed by taking all nodes within the respective regions and apply Equation (5.7). This gives rise to 8 distinct measures: $k_{\text{Niño1+2}}$, $l_{\text{Niño1+2}}$, $k_{\text{Niño3}}$, $l_{\text{Niño3}}$, $k_{\text{Niño4}}$, $l_{\text{Niño4}}$, $k_{\text{Niño3.4}}$, and $l_{\text{Niño3.4}}$. In order to reduce this vast amount of information, we do not make use of the (less frequently studied) Nino1+2 region and focus on the Nino3.4 region (which is the basis of the nowadays most common ONI index) and its two contributors, Nino3 and Nino4.

5.3.3.2 Volcano Regions

The locations of the three volcanoes responsible for the largest eruptions of the recent decades are shown in Figure 5.1. To obtain interpretable information on the (tele-)connectivity induced by these eruptions, we need to integrate the connectivity properties of a sufficiently large amount of meaningfully chosen grid cells. Therefore, we take the area within a radius of 5° around each volcano as basis for the regionalization procedure of k_i. This leads to the three observables k_{Pinatubo}, k_{Agung} and k_{Chichon}.

However, the aforementioned choice might not be optimal, since symmetric spatial regions in the near-field do not necessarily exhibit the strongest persistent temperature effects after an eruption. Instead, the specific local meteorological conditions (especially wind fields) during the eruption period largely control the three-dimensional patterns of atmospheric aerosol concentrations and, hence, the position of the strongest mid-term

cooling to be expected. Accordingly, the induced signal can be more evident within regions that have been shifted with respect to the locations of the volcanoes. To account for this, we also calculate regionalized degrees for accordingly shifted regions (see Section 5.4.2 for details), denoted as k'_{Pinatubo}, k'_{Agung} and k'_{Chichon}. Here, the specific regions have been selected according to an examination of the resulting degree fields succeeding the individual eruptions and the corresponding wind fields, seeking for the timing and position of the strongest anomalies in the degree field that could be attributed to each eruption. The resulting locations are depicted in Figures 5.5 (c), 5.5 (f) and 5.5 (i)). Note that although a volcanic eruption may start relatively abruptly, its larger-scale atmospheric effects commonly become effective only with a considerable delay of several months or more [McC1995; Rob2000].

5.4 Results & Discussion

The presentation of the results with the accompanying discussion is structured in two parts. First, we analyze the impacts of the El Niño–Southern Oscillation (ENSO) using the aforementioned global and localized observables. Afterwards, we will discuss the global SAT signatures of Mount Pinatubo's, Mount Agung's and El Chichon's eruptions.

5.4.1 El Niño–Southern Oscillation

We start with investigating the global effects of ENSO on the spatio-temporal co-variability structure of global SAT. From a complex network perspective, this problem has already been addressed in a variety of previous studies [Fan2017; Goz2008; Rad2013; Tso2008; Wie2016; Yam2008], making use of different approaches for constructing network structures from global climate data. However, none of these studies has considered the complementarity between topological and spatial network properties in detail, nor employed the concepts of modularity and global average link distance that constitute key aspects of this work.

Global Network Properties. The network transitivity $\mathcal{T}$ has been shown by Wiedermann et al. [Wie2016] to systematically discriminate between the EP and CP flavours of both types of ENSO events, El Niño and La Niña. While this reference used an area-weighted version of $\mathcal{T}$, we follow here the approach of Radebach et al. [Rad2013] using remapping onto an icosahedral grid. Hence, our area-weights are equal and the difference can be neglected. Figure 5.2 (a) shows the corresponding results obtained using our slightly modified data set. They are qualitatively almost indistinguishable from those of the two aforementioned studies as expected*.

To further quantify the strength of teleconnectivity in the global SAT field, the network modularity $\mathcal{Q}$ provides a prospective candidate measure that has not yet been

*Note, however, that Figure 5.2 shows the results for different network measures in dependence on the window midpoint, while Wiedermann et al. [Wie2016] used the endpoint, leading to a 6-month shift between the respective plots.

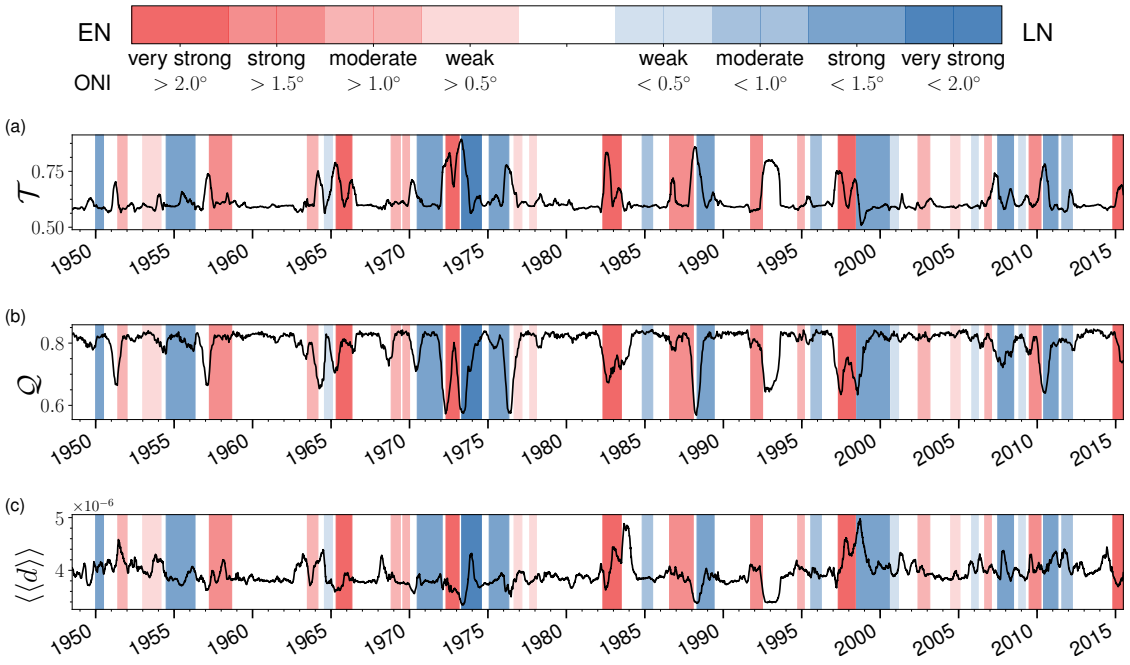


Figure 5.2: The time series of the observables (a) transitivity $\mathcal{T}$, (b) modularity $\mathcal{Q}$ and (c) global average link distance $\langle\langle l \rangle\rangle$. Background colors highlight different ENSO phases (red: El Niño (EN), blue: La Niña (LN)) according to the Ocean Niño Index (ONI), with opacity representing the corresponding strength. The ticks indicate January 1 of the corresponding year; all values are shown according to the midpoint dates of the respective time windows. These graphics were published in [Kit2017d].

applied for this purpose in previous studies. Recall that a high modularity indicates a fragmented network, whereas low values point to a relatively homogeneous connectivity structure of the network as a whole. Hence, a decrease in modularity could indicate an increase in the degree of organization of the global SAT network, i.e. a tendency towards more balanced co-variability in global temperatures.

Figures 5.2 (a) and 5.2 (b) show that most time intervals that are characterized by high values of network transitivity actually exhibit a reduction in modularity. Consistent with previous findings [Rad2013], most of these time windows in fact coincide with either a El Niño or La Niña phase. This indicates again the global impact of these episodes and can be considered as an expected signature of emerging teleconnectivity*.

*Note that taken alone, this process would not necessarily imply a stronger *synchronization* (as studied by, e.g. Maraun et al. [Mar2005]) between climate variability in distinct regions, which would be reflected by higher absolute correlation values. Specifically, in this work, we use a fixed link density of 0.5% in all window-specific climate networks and thus cannot make any statements about the overall strength of correlations. However, following previous results by Radebach et al. [Rad2013], we may actually expect that the correlation threshold $q_{|c|,0.995}$ used for establishing network connectivity in this work exhibits maxima whenever $\mathcal{T}$ shows a peak, thereby supporting the hypothesis of El Niño and La Niña episodes synchronizing global SAT variability by establishing teleconnections.

The temporal variability profiles of transitivity and modularity also exhibit some important differences. In particular, the strong 1982/83 El Niño is represented as a single long episode of reduced modularity values while being split into two rather distinct peaks in transitivity (see Figures 5.2 (a) and 5.2 (b)). Given the known seasonal profile of El Niño peaking around Christmas, it is remarkable that the ONI index has remained high during a quite long period of time. This indicates a single extended event even despite its temporary decay captured by $\mathcal{T}$, but not quite by $\mathcal{Q}$. This underlines that both measures actually capture different aspects of network organization that provide complementary information.

Another notable observation is related with the abrupt shift from El Niño to La Niña conditions in summer 1998, leading to a very fast reorganization of the global SAT field. The latter transition is reflected by some negative anomaly of $\mathcal{T}$ in summer 1998 in comparison with the "normal" background values of this measure. This presents a unique feature in the time evolution of network transitivity over the last decades that is not accompanied by any corresponding anomaly in $\mathcal{Q}$.

While network transitivity and modularity present two key topological network characteristics, functional climate networks are systems embedded in geographical space. Hence, the spatial placement of nodes and links (which is disregarded by topological characteristics) can play an important role in the network structure formation [Rad2013]. To address this aspect, we use the temporal evolution of the global average link distance $\langle\langle l \rangle\rangle$ in Figure 5.2 (c). This measure exhibits more irregular variability with a less clear distinction between "background level" and "anomalies" associated with different types of climate disruptions than the previously studied two topological characteristics. Yet, the general behavior of $\langle\langle d \rangle\rangle$ resembles transitivity in the sense that ENSO-related peaks often co-occur in both measures. This indicates that strong El Niño and La Niña episodes do not exclusively trigger short-range (localized) connectivity (high $\mathcal{T}$), but also global teleconnectivity (high $\langle\langle l \rangle\rangle$), which is in line with contemporary knowledge on the large-scale impacts of both types of ENSO phases.

From the results discussed above, we tentatively conclude that in order to distinguish globally influential ENSO events from episodes of minor (or more regional) relevance, a combination of modularity and average link distance can be useful, taking a holistic view in studying the differential imprints of different types of ENSO phases. We will recall this strategy when discussing the effects of volcanic eruptions on network organization at a global scale.

Spatial Patterns of Network Connectivity. To extend the discussion of global impacts of ENSO events on teleconnectivity, we have created the composite plots in Figure 5.3. A composite is the average of a field, here the degree k_i and the average link distance l_i, for different event types, i.e. EP El Niño, CP El Niño, EP La Niña, and CP La Niña (cf. Table 5.2). As a reference, we included a composite for all years that are not classified, too. As representatives for the field for each event, we use the values for Christmas (Dec 24) of the corresponding year, because this is the usual time where El Niño / La Niña have their peak intensity.

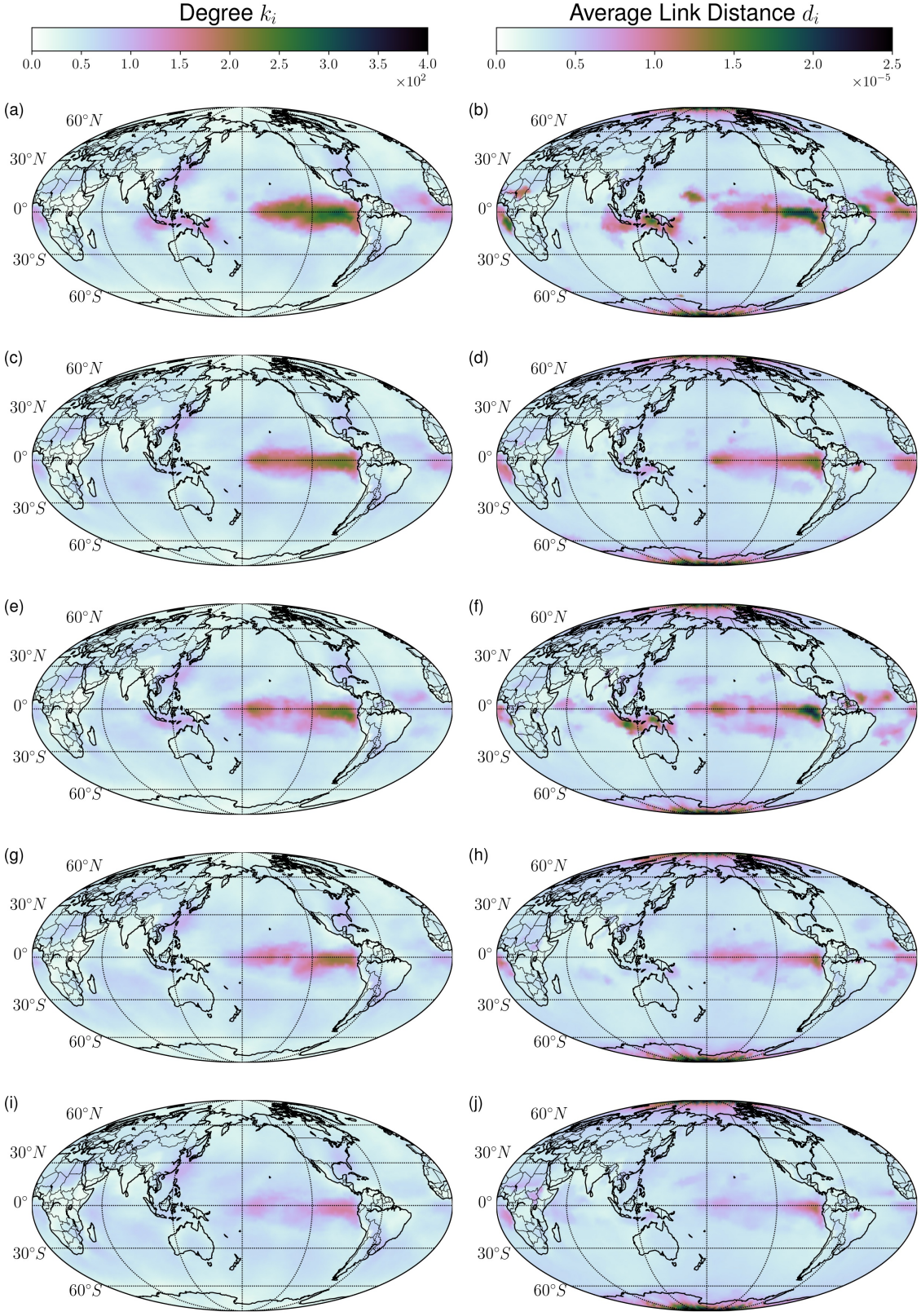


Figure 5.3: Global maps showing composites of (a,c,e,g,i) degree k_i and (b,d,f,h,j) average link distance l_i for different types of ENSO phases: (a,b) EP El Niño, (c,d) CP El Niño, (e,f) EP La Niña, (g,h) CP La Niña and (i,j) all other periods. The corresponding classification of different years is summarized in Table 5.2. These graphics were published in [Kit2017d].

The left panels of Figure 5.3 display the respective mean degree fields for the different types of ENSO periods. As expected, we observe a particularly strong deviation from a homogeneous pattern during EP El Niños (Figure 5.3 (a)), while the degrees in the Eastern-to-Central Tropical Pacific are only slightly larger than in the rest of the network during time windows without El Niño or La Niña conditions (Figure 5.3 (i)). Still, the observed degree patterns alone do not allow us to distinguish between a local or global phenomenon. For this purpose, the right panels of Figure 5.3 show the corresponding mean average link distance fields for each type of situation. High values of this measure in the typical ENSO region are present in case of all four possible types of episodes, indicating that both flavours of El Niño and La Niña actually generate additional connections in the Tropical Pacific that span relatively large distances.

Analyzing the composite maps of the average link distance in more detail, it is important to note that beyond the ENSO region itself, additional parts of the globe exhibit high values. This indicates the presence of localized teleconnections that possibly link climate variability in the latter regions with ENSO. Specifically, EP El Niños (Figure 5.3 (b)) exhibit such teleconnections with Indonesia and Western Africa, which are also recovered for EP La Niñas (Figure 5.3 (f)). For CP El Niños (Figure 5.3 (d)), the l_i field highlights a weak connection with Western Africa, but none with Indonesia. Similar but still weaker teleconnections can be observed for CP La Niñas (Figure 5.3 (h)).

The apparent teleconnection with Indonesia present during EP events, but not during their CP counterparts, is particularly interesting, as it is localized in the Westernmost Tropical Pacific. Thus, it connects Eastern and Western Pacific while not leading to marked long-distance connections in the central Pacific close to the dateline. It has to be noted that our analysis is based on cross-correlations only. The values can be severely affected by distinct temporal persistence properties of SAT in the Eastern and Western Tropical Pacific, as pointed out by recent studies making use of modern causal inference methods [Bal2013; Run2014]. From an impact perspective, the teleconnection suggested by our results is compatible with the documented increased likelihood of droughts in Indonesia during El Niño events [Dia2001].

The apparent teleconnection with Western Africa spans a rather large spatial distance (about one third of the globe). In this context, Joly et al. [Jol2009] noted that "a significant part of the West African monsoon interannual variability can be explained by the remote influence of ENSO." This previously reported teleconnection could be responsible for the elevated average link distance over Western Africa especially during EP El Niños.

Finally, the poles exhibit high values in all cases for the average link distance l_i. As these links are omnipresent, they are likely to be artifacts.

Regionalized Network Characteristics. The global and local climate network properties discussed above still provide only incomplete information on the effects of climate variability in different parts of the ENSO regions. Hence, we now analyze the regionalized field measures (see Section 5.3.3) and study the specific connectivity associated with the Nino3.4, Nino3 and Nino4 regions in terms of degree and average link distance.

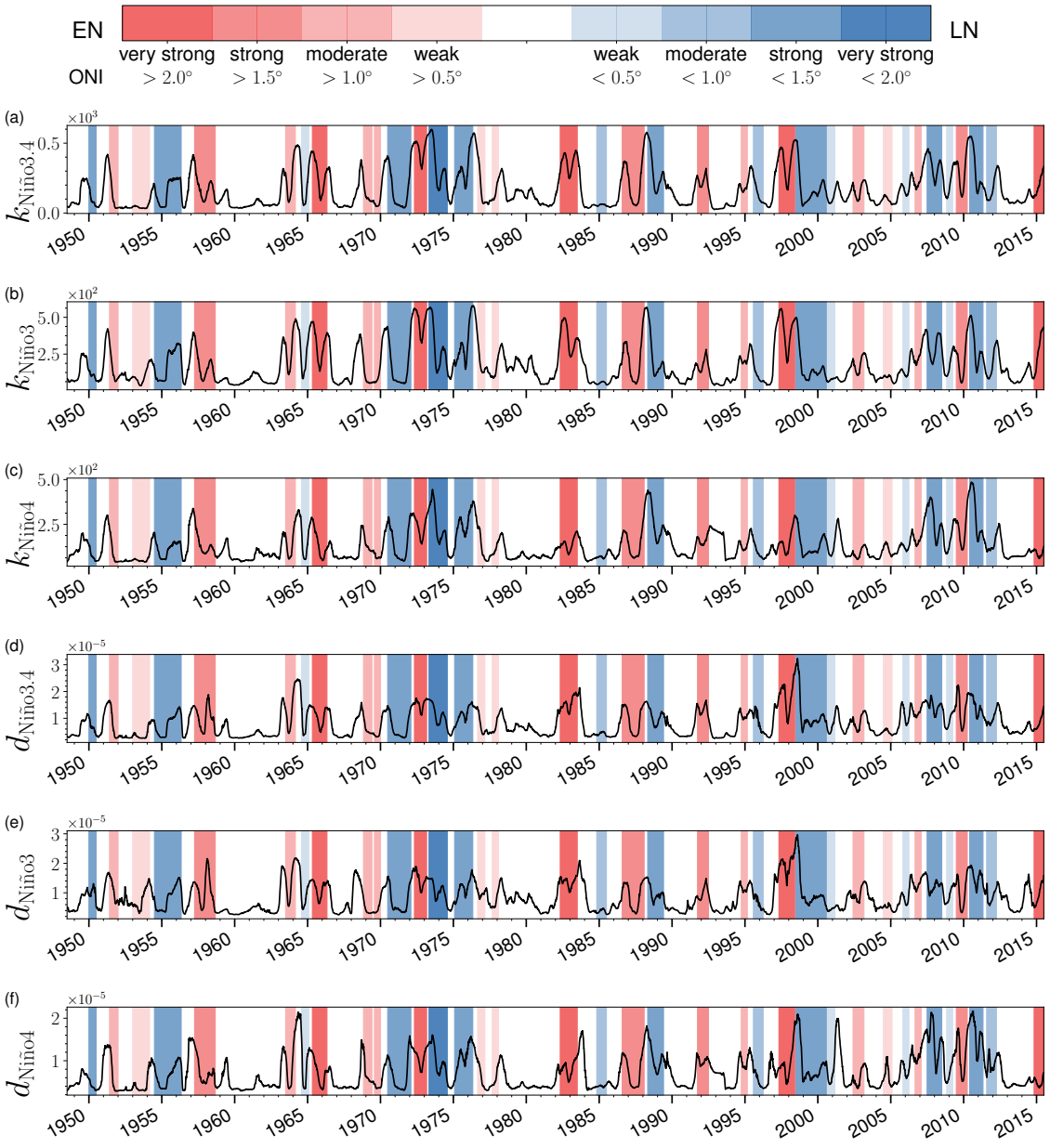


Figure 5.4: Time series of different regionalized climate network properties introduced in Section 5.3.3.1. Background colors and time axis as in Figure 5.2. These graphics were published in [Kit2017d].

The corresponding results are summarized in Figure 5.4. We observe that the relative magnitude of variations of regionalized degree and average link distance is even stronger than that of the global network properties transitivity, modularity and global average link distance discussed above. All measures exhibit episodes of very small values as opposed to such with much larger values, the latter often coinciding with El Niño and La Niña phases. Since the corresponding regions have been previously chosen for defining ENSO-specific indices, this result has been expected. Most importantly, degree and average link distance based characteristics exhibit strong positive correlations. For climatic events with predominantly local structure, we would expect a strong increase of k_i but only weaker increase of d_i in the respective region. Hence, our corresponding observations underline that ENSO-related climate impacts are not confined to the vicinity of the ENSO region, but are controlled by large-scale teleconnections.

Since the different ENSO regions show partial overlap (cf. Figure 5.1), the results obtained for the individual regions exhibit a high degree of similarity. However, regarding specific El Niño or La Niña episodes, comparing the corresponding signatures for the Nino3 and Nino4 regions still allows attributing these events to more Eastern Pacific or Central Pacific types. For example, the strong 1997/98 El Niño is reflected by very high values of the regionalized degree for the Nino3 and Nino3.4 regions, but relatively weak signatures in the more Western Nino4 region, which is consistent with its classification as an EP type event.

Next we examine the time evolution of all six regionalized network measures in some detail. It is notable that between 1978 and 1982, there has been considerable variability in all measures pointing towards an episodic presence of teleconnections even though none of the time windows was classified as an El Niño or La Niña episode according to the ONI. Moreover, we find that before the year 2000, clear peaks can always be observed in all properties as alternating with periods of low values. In contrast, during the last 15 years, we find strong variability with peaks occurring almost annually (except 2013 and 2014). This change in the overall temporal variability pattern of our regionalized network measures might point to some fundamental changes in the spatio-temporal organization of global SAT, either due to some not yet identified mode of natural variability or as a result of external interference.

5.4.2 Volcanoes

Besides ENSO variability, strong volcanic eruptions have been identified as causes of strong perturbations in climate network properties in earlier studies [Rad2013]. Hence, the application of the complementary viewpoints as used in this work for further characterizing the impacts of such eruptions promises interesting additional insights.

Regarding the global network properties, let us turn back to Figure 5.2. A signature similar to an ENSO event has been found in the aftermath of the Mount Pinatubo eruption in transitivity and modularity, suggesting that this event has affected the climate system globally. However, when comparing these topological network characteristics with the spatial network property of the global average link distance $\langle\langle l \rangle\rangle$, we find a distinct

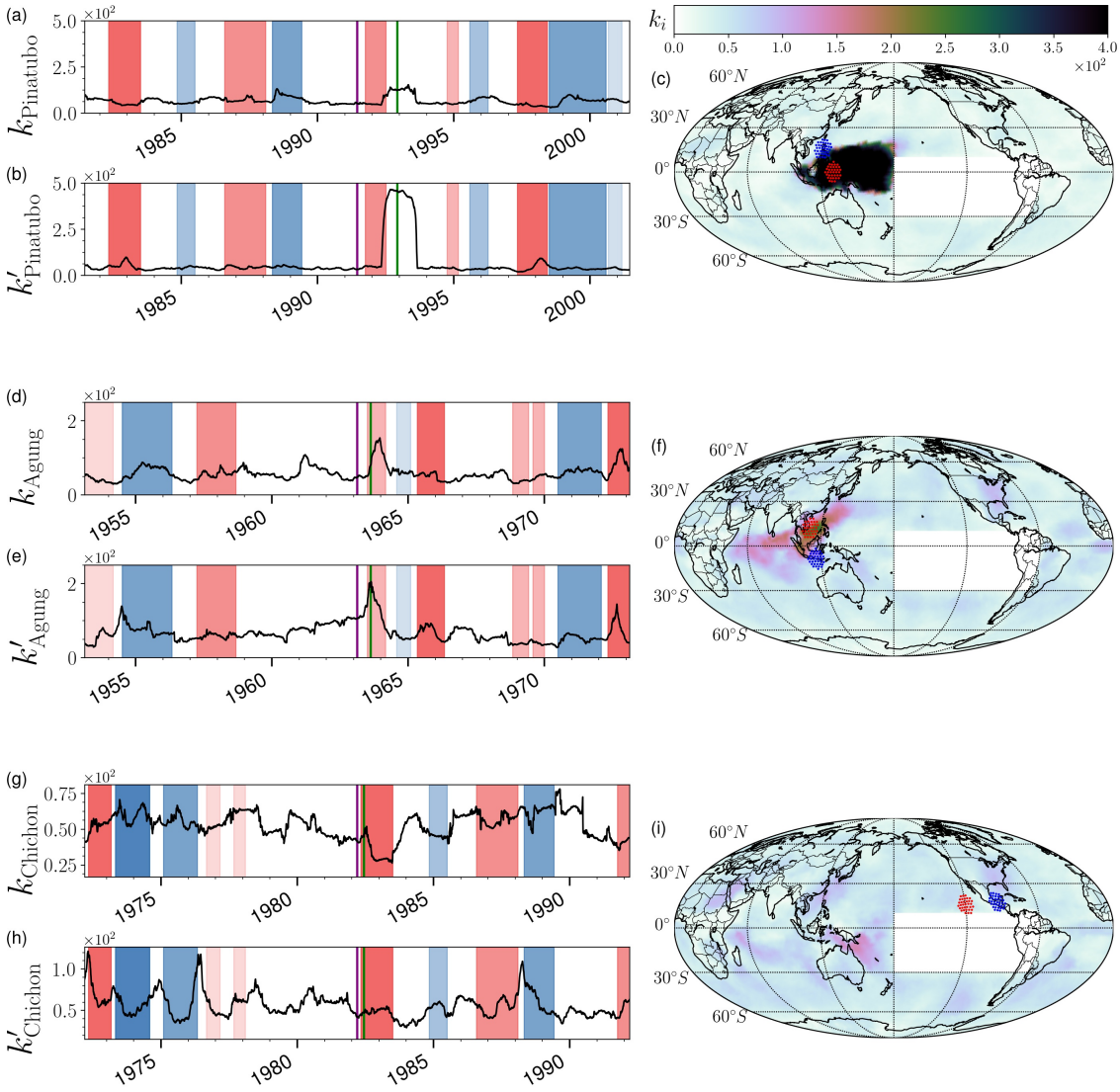


Figure 5.5: Time series of regional mean degree and associated degree field (excluding the ENSO-big region indicated by white patches in the corresponding maps) for the three main volcanic eruptions during the study period: (a-c) Mount Pinatubo, (d-f) Mount Agung, and (g-i) El Chichon. In the degree maps shown in panels (c), (f) and (i), blue dots mark grid points within a radius of 5° around each volcano, which have been used to define the regionalized degrees shown in panels (a), (d) and (g), respectively. Red dots indicate spatially shifted regions of the same size where the largest changes to the degree field have been observed. These regions serve as the basis for computing the regionalized degrees shown in panels (b), (e) and (h), respectively. Purple vertical lines indicate the timing of the respective eruptions, whereas green vertical lines indicate the midpoints of the time windows exhibiting the strongest signature in the regionalized network properties. The time series have been restricted to ±10 years around the date of the respective eruption. Background colors indicate the corresponding ENSO strength as in Figure 5.2. These graphics were published in [Kit2017d].

difference. Specifically, ENSO-related climate perturbations show peaks in both, $\langle\langle d\rangle\rangle$ and $\mathcal{T}$. In contrast, we observe a dip in $\langle\langle d\rangle\rangle$ with a corresponding peak in $\mathcal{T}$ in 1993, which corresponds to the time windows where the cooling effect following the Mount Pinatubo eruption should have taken its maximum [McC1995]. Hence, unlike for ENSO-related disruptions, this indicates that the effects of the volcanic eruptions have been mostly regionally confined. The latter is consistent with the hypothesis of elevated correlations in the region that has been most directly affected by the associated cooling trend following the eruption. Based on this observation, we suggest that using the global average link distance in conjunction with network transitivity and modularity enables us to discern disruptive events with global effects (strong ENSO phases) from those exhibiting more regional impacts (volcanic eruptions).

In general, large-scale effects of volcanic eruptions on global SAT teleconnectivity can be observed only after a sufficiently large amount of aerosols have entered the stratosphere [Rob2000]. Accompanying the resulting time shift between trigger event and response, we may also need to consider a spatial shift of the most affected region as compared to the location of the volcano. We have applied our regionalization procedure (see 5.3.3.2) to study the impacts of the Mount Pinatubo, Mount Agung and El Chichon eruptions. In order to avoid interference with the effects of ENSO events, the ENSO-big region depicted in Figure 5.1 are excluded from the corresponding computations. The results obtained from this analysis are in Figure 5.5.

The largest of the three considered eruptions (Mount Pinatubo) had a global cooling effect and has left clearly visible signatures in all considered global network measures as discussed above. Some months after the eruption, a large region of elevated network connectivity has established, which covers essentially all of the Western Tropical Pacific (Figure 5.5 (c)). The average degree in the region around Mount Pinatubo displays an abrupt rise about half a year after the eruption, then a constantly high value for about one year (the common residence time of volcanic aerosols in the stratosphere) before dropping again back to its previous level (Figure 5.5 (a)). The region with the highest degrees is shifted Northward with respect to the location of the volcano (Figure 5.5 (c)). When computing the average degree for this region, we observe an much stronger increase of the regionalized degree than for the region surrounding the volcano (Figure 5.5 (b)).

The Mount Agung eruption exhibits similar, but weaker, patterns in the respective region (Figure 5.5 (f)). However, the region with the highest degree is shifted South-Westward. The average degree in the region surrounding Mount Agung only shows weak changes after the eruption (Figure 5.5 (f)). In contrast, there is a sharper increase in the shifted region. The peak effect occurs significantly faster after the beginning of the eruption than in case of Mount Pinatubo (Figure 5.5 (e)).

Unlike the two other volcanic eruptions, the degree field in the period succeeding the El Chichon eruption showed hardly any marked changes (Figure 5.5 (i)). Consequently, we also do not observe any marked signature in the temporal variability profile of the regionalized degree in the surrounding of the volcano (Figure 5.5 (g)). Instead of a peak shortly after the eruption, we actually find a clear drop of the average degree. However, given that El Chichon is located relatively close to the extended ENSO region, we cannot

rule out that this could be an effect of the strong El Niño event occurring shortly after the eruption and eventually even being partially triggered by the latter [Kho2017].

5.5 Summary

In this chapter, we have used functional climate networks from correlations of global surface air temperature (SAT) data to analyze the global impact and teleconnections of past El Niño / La Niña events and volcanic eruptions. By making use of the global network property of modularity, we have found that at least the East Pacific flavours of such events lead to a *global reconfiguration* of SAT variations. Considering the global average link distance as a complementary spatial network characteristic, we have identified distinct qualitative differences between the imprints of these ENSO periods and the Mount Pinatubo eruption in global SAT patterns.

Using composites of the degree and average link distance fields, we have identified traces of distinct ENSO teleconnections in the climate network structure. Especially prominent were the ones linking the Eastern Tropical Pacific with Indonesia and West Africa during East Pacific El Niños. By making use of a regionalization procedure applied to these two fields of local network properties, we have introduced a simple yet effective tool to reveal the differential roles of different regions in the Tropical Pacific in establishing teleconnections during different El Niño and La Niña events.

Finally, we have analyzed the global and local connectivity properties of SAT-based climate networks after the strongest recent volcanic eruptions of Mount Pinatubo, Mount Agung and El Chichon. While the Mount Pinatubo eruption has been confirmed to exhibit marked impacts on global SAT, its dominating effect was rather regional (i.e. it did not trigger long-range teleconnections detectable by our approach).

In summary, our study confirms that ENSO does not only have a strong local effect on SAT in the Tropical Pacific with a spatially confined increase of network connectivity [Rad2013], but also dynamically reconfigures climate variability globally by triggering teleconnections especially with other Tropical regions.

Chapter 6

Conclusion

6.1 Transient Analysis of Complex Dynamical Systems

The aim of this thesis was to advance the techniques and methods for understanding the transient phase of a trajectory in a complex dynamical system in the context of sustainability. To be precise, I took my motivations from understanding transformations to a sustainable future while respecting certain boundaries. Hence, I have formulated two vital questions:

(i) "How can we properly quantify the time to reach a system's attractor?" and

(ii) "Can we avoid transgressing boundaries and stay safe (& just)?"

Subsequently, I have turned these two questions in two work packages *Timing of Transients* and *Topology of Sustainable Management* and addressed them within this thesis. My next step was to adapt these methods to time series analysis, but this did not turn out fruitful. Still, during the process I have found interesting signatures resulting in the third work package *Teleconnectivity Structures of the El Niño-Southern Oscillation.*

In the following, I conclude on my contributions to transient analysis in complex dynamical systems sorted by these three work packages.

6.2 Contribution of this Thesis and Outlook

6.2.1 Timing of Transients

I have analyzed multiple problems of common approaches for quantifying a reaching time and formulated four conditions a sensible metric should have. Next, I have proposed two metrics, *Regularized Reaching Time* (T_{RR}) and *Area under Distance Curve* (D), to answer question (i)*. To demonstrate their applicability, I analyzed multiple example models with increasing complexity: a linear system, a global carbon cycle, a generator in a power grid and finally the chaotic Rössler oscillator. Within the scope of *Timing on*

*I would like to acknowledge that Kevin Webster provided the proofs for the invariance of T_{RR} under smooth coordinate transformations.

Networks, I extended T_{RR} as a network metric yielding $T_{RR}^{\mathcal{N}}$. In particular, the presented definition of $T_{RR}^{\mathcal{N}}$ in this thesis is new in comparison to the one in the related publication [Mit2017b]. Again, this metric has been applied to multiple examples: a deterministic scale-free network of Rössler oscillators, an ensemble of random scale-free networks of Rössler oscillators, and the power grid of the United Kingdom. Within that analysis, relations to the topological features of the systems have been analyzed. I would like to emphasize that the numerical analysis of the networked models has been done by Chiranjit Mitra, who is the lead-author in [Mit2017b]. Finally, I have checked the proposed metrics against the aforementioned and gave an in-depth discussion how they solve question (i).

This discussion has demonstrated how the present work induces a need for future research. There are four directions, that I would like to point out specifically. (1) While the two properties convergence and smooth-invariance of T_{RR} could be proven for fixed-points, it is only a hypothesis for more complex attractors. The numerical estimations for the chaotic Rössler attractor have shown sensible results, still a definition that leads to a mathematical proof of the aforementioned properties should be targeted. Here, Lyapunov functions seem to be good candidates as I found T_{RR} and D to be such a function. (2) Continuing from (1) it seems natural to use more sophisticated estimation methods and algorithms from Lyapunov estimations in order to quantify T_{RR} and D. A starting point for this is the work by Giesl et al. [Gie2015]. (3) Relations between $T_{RR}^{\mathcal{N}}$ and the topological structure of the underlying networks have been identified, but more research is necessary to understand and explain them. (4) It seems natural to devise methods to estimate statistics of the presented metrics from climate data or artificial output from climate models.

Related Publications:

- [Kit2017f] **T. Kittel**, J. Heitzig, K. N. Webster, and J. Kurths. "Timing of transients: Quantifying reaching times and transient behavior in complex systems". In: *New Journal of Physics* (2017). URL: `http://iopscience.iop.org/10.1088/1367-2630/aa7b61`
- [Mit2017b] C. Mitra, **T. Kittel**, A. Choudhary, J. Kurths, and R. V. Donner. "Recovery time after localized perturbations in complex dynamical networks". In: *New Journal of Physics* 19.10 (2017), p. 103004. URL: `http://stacks.iop.org/1367-2630/19/i=10/a=103004`

6.2.2 Topology of Sustainable Management

Under the name *Topology of Sustainable Management* (TSM), Jobst Heitzig developed a *mathematical framework* to analyze models with respect to boundaries in state space [Hei2016]. The latter are incorporated by the distinction between "desirable and undesirable states". Furthermore, the framework considers "management", i.e. a (possibly small) control of the system, and distinguishes it from the "default dynamics".

Starting from the original mathematical framework of TSM I have devised a variant definition of this framework based on concepts of *Viability Theory* (VT). Furthermore, I

extended the definition by introducing the important distinction between so-called time-limited and time-unlimited lakes*. This connection to VT has enabled me to use its tools. In particular I have adjusted the Saint-Pierre algorithm (SPA) so one can apply it to TSM, calling this step the *operationalization of TSM*. Within the variant definition of TSM the so-called eddies are defined by implicit capture basins. For that reason, I have extended the SPA to deal with these kind of sets, too. During the application to the example model, two common problems of estimations using SPA arose: an unbounded state space and highly varying time scales. I have solved both by introducing appropriate, general coordinate transformations, a *compactification of the state space* and the *time homogenization*.

The aforementioned concepts were applied to multiple, two-dimensional models which I have analyzed manually. Taking the step to higher dimensional models induced the need for an automatic analysis which I have done with the aforementioned operationalization. To demonstrate this, I have applied the SPA including my extension for implicit capture basins, the compactification of the state space and the time homogenization to a three-dimensional model focusing on climate change, economic output and energy transformation, called the AYS-model†. I have developed the idea of TSM-bifurcations as qualitative changes of the TSM-topology structure. Within this model, I have demonstrated a first TSM-bifurcation analysis and found a so-called "downstream-eddies bifurcation".

Within the concept of TSM dilemmas have been identified, most notably the *lake dilemma*. This is a situation with a qualitative choice between (a) uninterrupted desirability (that is, not transgressing the boundaries) or (b) transgressing the boundaries temporarily but finally reaching a region where one does not have to apply management anymore. With the distinction between the aforementioned time-limited and time-unlimited lake introduced in this thesis, an aggravated form of the lake dilemma has been identified. The *pressing lake dilemma* that happens inside the time-limited lake demands for the decision between (a) and (b) within a limited time frame. Note that not actively taking a decision is still a decision.

In the process of this work, I published a basic software library for TSM estimations called `PyViability`‡ and I published the Code for the reproduction of the analysis results of the 3-dimensional example model§.

With this work I hope to have taken some first steps to sharpen the debate on transitions to a sustainable future. The example system was of low complexity so I could focus on the operationalization of TSM. For that reason, relevant parts, e.g. a carbon cycle, an economic cycle and different energy productions, were represented simplistically only.

*Within this conclusion, lakes can be see as situations where the "lake dilemma", explained further down in the text, occurs. A time-limited lake is then the situation where the lake dilemma is aggravated to the pressing lake dilemma (also further down in the text). In contrast, the time-unlimited lake refers to situations where the lake dilemma, but not the pressing lake dilemma, occurs. If the reader is interested in a precise definition of a lake and how the mentioned dilemmas arise, I have to refer him to Sections 4.2 and 4.3.

†This model has been contributed by Jobst Heitzig and I did the analysis.

‡`https://timkittel.github.io/PyViability/`

§`https://github.com/timkittel/ays-model`

This induces the need for more complex models, where a candidate seems to be the one in [Nit2016; Nit2017]. The complexity is higher than the one used in this article but the analysis should still be possible.

Because social foundations are often defined on distributions [Raw2012], e.g. the income distribution, another branch of research would be the development of modeling techniques and model reductions for said distributions. Network-based approaches with thermodynamical limits or pair-approximations seem plausible, e.g. [Wie2015].

With an increased dimensionality of the system, there is a need for improved algorithms also. Here, existent algorithms are found in [Alv2016; Bok2006; Bri2016; Mai2013], but they need to be adjusted and extended to fit to the computation of the TSM partition.

Related Publications:

- [Hei2016] J. Heitzig, **T. Kittel**, J. F. Donges, and N. Molkenthin. "Topology of sustainable management of dynamical systems with desirable states: from defining planetary boundaries to safe operating spaces in the Earth system". In: *Earth System Dynamics* 7.1 (2016), pp. 21– 50. doi: 10.5194/esd-7-21-2016. URL: https://www.Earth-syst-dynam.net/7/21/2016/
- [Kit2017e] **T. Kittel**, J. Heitzig, G. Deffuant, J.-D. Mathias, and J. Kurths. "Operationalization of Topology of Sustainable Management to Estimate Qualitatively Different Regions in State Space". 2017. arXiv: 1706.04542. URL: https://arxiv.org/abs/1706.04542 (submitted to Environmental Modelling & Software)

Related Code Publications:

- [Kit2017c] **T. Kittel**. "The PyViability Library". 2017. URL: https://timkittel.github.io/PyViability/
- [Kit2017b] **T. Kittel**. "Code for reproduction of the TSM analysis for the AYS model". 2017. URL: https://github.com/timkittel/ays-model/

6.2.3 Teleconnectivity Structures of the El Niño-Southern Oscillation

Based on the approach of functional climate networks constructed from spatial correlations of daily global surface air temperature (SATI have presented), I have analyzed the global impact and teleconnections of past El Niño and La Niña events. Furthermore, I applied this analysis to the three major volcanic eruptions since the middle of the 20^{th} century, Mount Pinatubo (1991), Mount Agung (1963) and El Chichon (1982).

Using the global measure modularity, I have found that the Eastern Pacific flavors of the ENSO (EP El Niño and EP La Niña) lead to a global reconfiguration of the SAT variations. As a complementary spatial network characteristic, the global average link distance, enabled me to identify distinct qualitative differences between ENSO periods and Mount Pinatubo.

Using the composites of degree and average link distance, I have identified traces of distinct teleconnections of the ENSO region, especially to Indonesia and West Africa

during EP El Niños. With the introduced regionalization procedure, I found a simple, but still effective tool to differentiate which roles different tropical Pacific regions have in establishing teleconnections during different El Niño and La Niña phases.

Finally, I have analyzed the global and local properties of functional climate networks based on SAT after the aforementioned volcanic eruptions. While Mount Pinatubo has a distinct and strong impact on the global SAT, its dominating effect was very regional, i.e. there were no long-range teleconnections detected.

A natural next step is to take the results of this analysis and make rigorous statistical significance test on the proposed teleconnections in order to uncover the underlying climatological processes. A different direction of research would be to systematically analyze and understand the remapping algorithm, proposed in [Rad2013] and used here to discern possible artifacts from actual signals.

Related Publications:

- [Kit2017d] **T. Kittel**, C. Ciemer, N. Lotfi, T. Peron, F. Rodrigues, J. Kurths, and R. V. Donner. "Global teleconnectivity structures of the El Niño–Southern Oscillation and large volcanic eruptions: An evolving network perspective". 2017. (submitted to Nonlinear Processes in Geophysics)

Related Code Publications:

- [Kit2017a] **T. Kittel**. "Code for reproduction of global teleconnectivity structures of ENSO and large volcanic eruptions". 2017. URL: `https://github.com/timkittel/global-structures-enso-volcanoes/`

Appendix A

Regularized Reaching Time for Hyperbolic Fixed Points

This appendix is taken from [Kit2017f].

We consider a deterministic dynamical system of the form

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^n, \quad f \in C^2(\mathbb{R}^n, \mathbb{R}^n) \tag{A.1}$$

and assume that the system contains an exponentially asymptotically stable equilibrium $\boldsymbol{x}^{\mathrm{f}}$ at the origin, hence the corresponding attractor set is $\mathscr{A} = \{\boldsymbol{x}^{\mathrm{f}}\}$. That is, $f(\boldsymbol{x}^{\mathrm{f}}) = 0$ and all the eigenvalues of $Df(\boldsymbol{x}^{\mathrm{f}})$ have negative real part, where D is the differentiation operator. We denote the basin of attraction of $\boldsymbol{x}^{\mathrm{f}}$ by $\mathscr{B}_{\mathscr{A}}$. Also we denote the flow operator of (A.1) as $\varphi(t, \cdot)$.

Let the spectrum of $Df(\boldsymbol{x}^{\mathrm{f}})$ be $\{\lambda^s\} \cup \sigma^{ss}$, where λ^s may be real or complex, but we assume has multiplicity one. It will also be useful later to define constants α^s and α^{ss} such that

$$\{\sigma^{ss}\} < \alpha^{ss} < \mathrm{Re}\ \lambda^s < \alpha^s < 0, \tag{A.2}$$

where we also require $2|\alpha^s| > |\mathrm{Re}\ \lambda^s|$.

In the following, we further assume that the function $\Delta(\cdot)$ from Equation (3.7) has been chosen such that $\Delta(x) = ||\boldsymbol{x} - \boldsymbol{x}^{\mathrm{f}}||_{\mathcal{N}} = ||\boldsymbol{x}||_{\mathcal{N}}$ for some norm $||\cdot||_{\mathcal{N}}$ where the second step is possible as we chose $\boldsymbol{x}^{\mathrm{f}}$ to be at the origin. If $\boldsymbol{x}^{\mathrm{f}}$ were not at the origin, one could shift the coordinates such that in the new coordinate system it is. The conditions on $||\cdot||_{\mathcal{N}}$ are given in Proposition A.4. So we can restate the t^ϵ from Equation (3.7) as

$$t^\epsilon(\boldsymbol{x}) := \inf\{T : ||\varphi(t, \boldsymbol{x})||_{\mathcal{N}} < \epsilon \text{ for all } t \geq T\}. \tag{A.3}$$

The T_{RR} function is then defined as follows.

Definition A.1 (T_{RR} function) *For a given reference point $\boldsymbol{x}^{ref} \in \mathscr{B}_{\mathscr{A}}$, the T_{RR} function of a point $\boldsymbol{x} \in \mathscr{B}_{\mathscr{A}}$ is defined as*

$$T_{RR}(\boldsymbol{x}; \boldsymbol{x}^{ref}) := \lim_{\epsilon \to 0} \left[t^\epsilon(\boldsymbol{x}) - t^\epsilon(\boldsymbol{x}^{ref}) \right], \tag{A.4}$$

if the limit exists.

A natural question is under what conditions the limit in (A.4) exists. To answer this question we distinguish between two cases according to whether λ^s is real or complex. Our first result is regarding the existence of the T_{RR} function in both cases, and the dependence on the choice of norm $||\cdot||_{\mathcal{N}}$. In order to state the result for λ^s complex, we make the following definition.

Definition A.2 *Consider the system* (A.1) *and let* $\mathbb{R}^d = E^s \oplus E^{ss}$ *be the direct sum decomposition into the invariant stable and strong stable subspaces for the linear system* $\dot{\boldsymbol{x}} = Df(\boldsymbol{x}^f)\boldsymbol{x}$. *Let* $W^s(\boldsymbol{x}^f)$ *and* $W^{ss}(\boldsymbol{x}^f)$ *be the corresponding stable and strong stable manifolds of* $\boldsymbol{x}^f$.

Definition A.3 *We define the following equivalence class on norms defined on* $\mathscr{B}_{\mathscr{A}}$*:*

$$||\cdot||_{\mathcal{N}_1} \sim ||\cdot||_{\mathcal{N}_2} \quad \Leftrightarrow \quad ||v||_{\mathcal{N}_1} = ||v||_{\mathcal{N}_2} \text{ for all } v \in E^s. \tag{A.5}$$

Clearly, elements in the above equivalence class are defined by the norm of elements in E^s.

Proposition A.4 *Let the* T_{RR} *function be defined as in Definition A.1 for the system* (A.1)*, and assume* $\boldsymbol{x} \notin W^{ss}(\boldsymbol{x}^f)$. *Then we have the following*

1. *When* λ^s *is real, the limit* (A.4) *exists for all choices of norm* $||\cdot||_{\mathcal{N}}$. *Moreover, the limit is independent of the choice of norm.*
2. *When* λ^s *is complex, the limit* (A.4) *exists if and only if* $||\cdot||_{\mathcal{N}} \sim ||\cdot||_P$, *where* $||\boldsymbol{x}||_P := ||P^{-1}\boldsymbol{x}||_2$, $||\cdot||_2$ *is the Euclidean 2-norm, and* $P^{-1}Df(\boldsymbol{x}^f)P$ *is the Jordan normal form of* $Df(\boldsymbol{x}^f)$.

We will also show that the T_{RR} function is closely related to the strong stable foliation $\mathcal{F}^{ss}$ in the basin of attraction of the equilibrium $\boldsymbol{x}^{\mathrm{f}}$. We first recall the following definitions.

Definition A.5 *A foliation* $\mathcal{F}$ *of an* n*-dimensional manifold* M *is a partition of* M *into a disjoint collection of* k*-dimensional injectively immersed connected submanifolds (called leaves) such that for each* $\boldsymbol{x} \in M$, *there is a neighborhood* $V \subset M$ *and a chart*

$$\phi : V \to \mathbb{R}^k \times \mathbb{R}^{d-k}, \tag{A.6}$$

such that each connected component of the intersection of a leaf of $\mathcal{F}$ *with* V *is mapped to the set* $\mathbb{R}^k \times \{\boldsymbol{y}\}$, *for some* $\boldsymbol{y} \in \mathbb{R}^{n-k}$.

We call $\mathcal{F}$ *a* C^r *foliation if each local chart is* C^r. *A continuous foliation whose leaves are* C^r *is called a* C^r *lamination.*

Appendix A. Regularized Reaching Time for Hyperbolic Fixed Points

We denote the leaf of a foliation through a point $x \in \mathcal{B}_{\mathcal{A}}$ as $\mathcal{F}(x)$. A foliation $\mathcal{F}$ is invariant under the flow of (A.1) if $\varphi(t, \mathcal{F}(x)) = \mathcal{F}(\varphi(t, x))$ for sufficiently small $|t|$.

Theorem A.6 ([Hir1977]) *Consider the system* (A.1) *and E^s, E^{ss} and W^{ss} as in Definition A.2. Then there exists a unique invariant C^r lamination $\mathcal{F}_{ss}$ in $\mathcal{B}_{\mathcal{A}}$, called strong stable foliation, such that each leaf of $\mathcal{F}_{ss}$ has dimension equal to dim E^s and $\mathcal{F}_{ss}(x^f) = W^{ss}(x^f)$.*

Solutions $\varphi(t, x)$ and $\varphi(t, y)$ for initial conditions x, $y \in \mathcal{B}_{\mathcal{A}}$ that belong to the same leaf of $\mathcal{F}_{ss}$ for all time are characterized by strong asymptotic convergence to each other: $||\varphi(t, x) - \varphi(t, y)||_{\mathcal{N}} \leq Ce^{-\alpha^{ss}t}$ for t sufficiently large.

We will also prove the following result which provides an important characterization of the T_{RR} function.

Proposition A.7 *Let the T_{RR} function be defined as in Definition A.1 for the system* (A.1)*, and assume $x^{ref} \notin W^{ss}(x^f)$. In the case λ^s is complex, we assume $||\cdot||_{\mathcal{N}} \sim ||\cdot||_P$ as in Proposition A.4. Then the level sets of $T_{RR}(x; x^{ref})$ are equal to the leaves of $\mathcal{F}_{ss}$. That is,*

$$T_{RR}(x; x^{ref}) = T_{RR}(y; x^{ref}) \Leftrightarrow \mathcal{F}_{ss}(x) = \mathcal{F}_{ss}(y) \tag{A.7}$$

Furthermore, $T_{RR}(x; x^{ref}) \to -\infty$ as x approaches $\mathcal{F}_{ss}(x^f)$ $(= W^{ss}(x^f))$.

The proof of Propositions A.4 and A.7 rely on the following result regarding the behavior of solutions in the approach to equilibrium. We refer to [San1993] for a proof.

Theorem A.8 *Consider the system* (A.1)*, and define λ^s, α^s and α^{ss} as before. Then there exists $\kappa > 0$ such that for $x \in \mathcal{B}_{\mathcal{A}}$ with $||x - x^f||_{\mathcal{N}} < \kappa$, the limit*

$$\eta(x) := \lim_{t \to \infty} \Phi(0, t) P^s(\varphi(t, x) - x^f) \tag{A.8}$$

exists, where $\Phi(t, 0)$ is the transition matrix of $\dot{x} = Df(x^f)x$ from 0 to t and P^s is the projection onto E^s along E^{ss}. Furthermore, we have the representation

$$\varphi(t, x) - x^f = \Phi(t, 0)\eta(x) + \mathcal{O}(e^{-\min\{|\alpha^{ss}|, 2|\alpha^s|\}t}). \tag{A.9}$$

Note that since $\Phi(0, t)$ leaves E^s invariant and E^s is closed, we have $\eta(x) \in E^s$. It also follows from the proof in [San1993] that $\eta : B_\kappa(x^{\mathrm{f}}) \to E^s$ is continuous.

Lemma A.9 *Let $\varphi(t, x), \varphi(t, y)$ be solutions to* (A.1) *for $x, y \in \mathcal{B}_{\mathcal{A}}$. Then $\varphi(t, x), \varphi(t, y)$ belong to the same leaf of $\mathcal{F}_{ss}$ for all t if and only if $\eta(\varphi(s, x)) = \eta(\varphi(s, y))$ for s sufficiently large. Furthermore, $\eta(\varphi(s, x)) = 0$ if and only if $\varphi(t, x) \in W^{ss}(x^f)$.*

Proof. By Theorem A.6, the solutions $\varphi(t, x)$ and $\varphi(t, y)$ belong to the same leaf of $\mathcal{F}_{ss}$ for all $t \geq 0$ if and only if $||\varphi(t, x) - \varphi(t, y)||_{\mathcal{N}} \leq Ce^{-\alpha^{ss}t}$. Let $s > 0$ be large

enough so that $||\varphi(s,\boldsymbol{x})-\boldsymbol{x}^{\mathrm{f}}||_{\mathcal{N}}, ||\varphi(s,\boldsymbol{y})-\boldsymbol{x}^{\mathrm{f}}||_{\mathcal{N}}<\kappa$. Now from Theorem A.8 we have $\eta(\varphi(s,\boldsymbol{x}))$, $\eta(\varphi(s,\boldsymbol{y}))\in E^s$, and Eq. (A.9) implies that this is possible if and only if $\eta(\varphi(s,\boldsymbol{x}))=\eta(\varphi(s,\boldsymbol{y}))$. The last statement follows directly from (A.9) and the theory of stable/unstable manifolds. □

Proof of Proposition A.4. Let $\boldsymbol{x}$ and $\boldsymbol{x}^{\mathrm{ref}}$ be as in Definition A.1 and let $\varphi(t,\boldsymbol{x})$, $\varphi(t,\boldsymbol{x}^{\mathrm{ref}})$ be the solutions to (A.1). Let $s>0$ be large enough so that $||\varphi(s,\boldsymbol{x})-\boldsymbol{x}^{\mathrm{f}}||_{\mathcal{N}}$, $||\varphi(s,\boldsymbol{x}^{\mathrm{ref}})-\boldsymbol{x}^{\mathrm{f}}||_{\mathcal{N}}<\kappa$, and let $\epsilon>0$ be small. Then from Theorem A.8, the solution $\varphi(t,\boldsymbol{x})$ intersects the boundary of the D-ball $B_\epsilon(\boldsymbol{x}^{\mathrm{f}})$ when

$$||\Phi(t,s)\eta(\varphi(s,\boldsymbol{x}))+\mathcal{O}(e^{-\min\{|\alpha^{ss}|,2|\alpha^{s}|\}t})||_{\mathcal{N}}=\epsilon \tag{A.10}$$

We first consider the case where λ^s is real. Then $\eta(\varphi(s,x))\in E^s$ is one dimensional and we obtain

$$t-\frac{1}{\lambda^s}\ln\epsilon = -\frac{1}{\lambda^s}\ln||\eta(\varphi(s,\boldsymbol{x}))+g(t)||_{\mathcal{N}}=:F_{re}(t), \tag{A.11}$$

$$\hat{t}-\frac{1}{\lambda^s}\ln\epsilon = -\frac{1}{\lambda^s}\ln||\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}}))+\hat{g}(\hat{t})||_{\mathcal{N}}=:\hat{F}_{re}(\hat{t}), \tag{A.12}$$

where $g(t),\hat{g}(t)=\mathcal{O}(e^{-\delta t})$ for some $\delta>0$. Since the solution $\varphi(t,\boldsymbol{x})$ and its derivative are bounded in $B_\kappa(\boldsymbol{x}^{\mathrm{f}})$, the same is true for $g(t)$, $\hat{g}(t)$. Therefore, for $\epsilon>0$ sufficiently small, Eqs. (A.11) and (A.12) can be uniquely solved for $t(\epsilon)=t$ and $\hat{t}(\epsilon)=\hat{t}$ respectively, and $t(\epsilon),\hat{t}(\epsilon)\to\infty$ as $\epsilon\to 0$. We refer to Figure A.1 (a) for a sketch of the solutions to Eqs. (A.11) and (A.12). Then $t^\epsilon(\boldsymbol{x})=t(\epsilon)$ and $t^\epsilon(\boldsymbol{x}^{\mathrm{ref}})=\hat{t}(\epsilon)$ and we have

$$t^\epsilon(\boldsymbol{x})-t^\epsilon(\boldsymbol{x}^{\mathrm{ref}}) = \frac{1}{\lambda^s}\ln\frac{||\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}}))-\hat{g}(\hat{t}(\epsilon))||_{\mathcal{N}}}{||\eta(\varphi(s,\boldsymbol{x}))-g(t(\epsilon))||_{\mathcal{N}}} \tag{A.13}$$

$$\lim_{\epsilon\to 0}\left[t^\epsilon(\boldsymbol{x})-t^\epsilon(\boldsymbol{x}^{\mathrm{ref}})\right] = \frac{1}{\lambda^s}\ln\frac{||\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}}))||_{\mathcal{N}}}{||\eta(\varphi(s,\boldsymbol{x}))||_{\mathcal{N}}}. \tag{A.14}$$

Recall $\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}}))$, $\eta(\varphi(s,\boldsymbol{x}))\in E^s$ and E^s is one dimensional, so the above is independent of the norm and we have

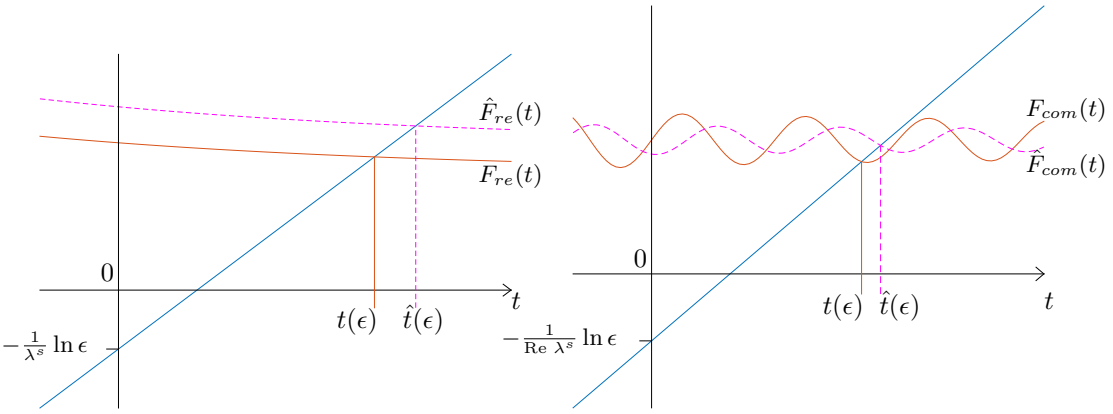
$$T_{RR}(\boldsymbol{x};\boldsymbol{x}^{\mathrm{ref}})=\frac{C(\boldsymbol{x})}{\lambda^s} \tag{A.15}$$

for some $C(\boldsymbol{x})\in\mathbb{R}$. This proves the first part of Proposition A.4.

Now assume that λ^s is complex. Then Eq. (A.10) leads to

$$\begin{aligned} t-\frac{1}{\mathrm{Re}\ \lambda^s}\ln\epsilon &= -\frac{1}{\mathrm{Re}\ \lambda^s}\ln\left|\left|P\Psi(t,s)P^{-1}\eta(\varphi(s,\boldsymbol{x}))+g(t)\right|\right|_{\mathcal{N}} \\ &=: F_{com}(t), \end{aligned} \tag{A.16}$$

$$\begin{aligned} \hat{t}-\frac{1}{\mathrm{Re}\ \lambda^s}\ln\epsilon &= -\frac{1}{\mathrm{Re}\ \lambda^s}\ln\left|\left|P\Psi(\hat{t},s)P^{-1}\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}}))+\hat{g}(\hat{t})\right|\right|_{\mathcal{N}} \\ &=: \hat{F}_{com}(\hat{t}), \end{aligned} \tag{A.17}$$



(a) Schematic diagram illustrating solutions of (A.11) and (A.12).

(b) Schematic diagram illustrating solutions of (A.16) and (A.17) for a choice of $||\cdot||_{\mathcal{N}}$ not in the equivalence class of $||\cdot||_P$, so that $t(\epsilon) - \hat{t}(\epsilon)$ oscillates as $\epsilon \to 0$ (see text).

Figure A.1: Schematic diagrams illustrating the different kinds of solutions. These graphics were published in the supplementary data of [Kit2017f].

where $P^{-1}Df(\boldsymbol{x}^{\mathrm{f}})P$ is the Jordan normal form of $Df(\boldsymbol{x}^{\mathrm{f}})$, and $e^{\mu t}\Psi(t,s)$ is the transition matrix of $\dot{w} = P^{-1}Df(\boldsymbol{x}^{\mathrm{f}})Pw$ from s to t. Again $g(t), \hat{g}(t) = \mathcal{O}(e^{-\delta t})$ for some $\delta > 0$. However now the term $P\Psi(t,s)P^{-1}\eta(\varphi(s,\boldsymbol{x}))$ oscillates as $t \to \infty$, although it will remain bounded. Then for $\epsilon > 0$ small enough, the above equation will have a (not necessarily unique) solution for t. We define $t(\epsilon)$ and $\hat{t}(\epsilon)$ as the largest such solutions for given ϵ respectively, then again $t(\epsilon)$, $\hat{t}(\epsilon) \to \infty$ as $\epsilon \to 0$ and $t^\epsilon(\boldsymbol{x}) = t(\epsilon)$, $t^\epsilon(\boldsymbol{x}^{\mathrm{ref}}) = \hat{t}(\epsilon)$. We refer to Figure A.1 (b) for a sketch of the solutions to Eqs. (A.16) and (A.17). Then we have

$$\begin{aligned} &t^\epsilon(\boldsymbol{x}) - t^\epsilon(\boldsymbol{x}^{\mathrm{ref}}) \\ &\quad = \frac{1}{\mathrm{Re}\ \lambda^s} \ln \frac{\left|\left|P\Psi(\hat{t}(\epsilon),s)P^{-1}\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}})) + \hat{g}(\hat{t}(\epsilon))\right|\right|_{\mathcal{N}}}{||P\Psi(t(\epsilon),s)P^{-1}\eta(\varphi(s,\boldsymbol{x})) + g(t(\epsilon))||_{\mathcal{N}}} \end{aligned} \tag{A.18}$$

Note that

$$||\Psi(t,s)P^{-1}\eta(\varphi(s,\boldsymbol{x}))||_2 = ||P^{-1}\eta(\varphi(s,\boldsymbol{x}))||_2 \tag{A.19}$$

since $\eta(\varphi(s,\boldsymbol{x})) \in E^s$. Now if $||\cdot||_{\mathcal{N}} \sim ||\cdot||_P$ as in Proposition A.4, then we have

$$\begin{aligned} &\lim_{\epsilon\to 0}\left[t^\epsilon(\boldsymbol{x}) - t^\epsilon(\boldsymbol{x}^{\mathrm{ref}})\right] \\ &\qquad = \frac{1}{\mathrm{Re}\ \lambda^s} \ln \frac{||P^{-1}\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}}))||_2}{||P^{-1}\eta(\varphi(s,\boldsymbol{x}))||_2} \end{aligned} \tag{A.20}$$

so the T_{RR} function is well defined. However, for any other choice of norm $||\cdot||_{\mathcal{N}}$, the quantity $||P\Psi(t(\epsilon),s)P^{-1}\eta(\varphi(s,\boldsymbol{x}))||_{\mathcal{N}}$ will not converge to a constant value, and will oscillate as $t(\epsilon) \to \infty$. In particular, for any $\eta(\varphi(s,\boldsymbol{x}))$ with $\eta(\varphi(s,\boldsymbol{x})) \neq C\eta(\varphi(s,\boldsymbol{x}^{\mathrm{ref}}))$, the function $t^D(\boldsymbol{x},\epsilon) - t^D(\boldsymbol{x}^{\mathrm{ref}},\epsilon)$ will not converge as $\epsilon \to 0$. This proves the second part of Proposition A.4. □

Proof of Proposition A.7. From Lemma A.9, we have that

$$\boldsymbol{x},\, \boldsymbol{y} \in \mathcal{F}_{ss}(\boldsymbol{x}) \quad \Leftrightarrow \quad \eta(\varphi(s,\boldsymbol{x})) = \eta(\varphi(s,\boldsymbol{y})) \tag{A.21}$$

for s sufficiently large. From (A.14) and (A.20) we have

$$T_{RR}(\boldsymbol{x};\boldsymbol{x}^{\mathrm{ref}}) - T_{RR}(\boldsymbol{y};\boldsymbol{x}^{\mathrm{ref}}) = \frac{1}{\mathrm{Re}\ \lambda^s} \ln \frac{||\eta(\varphi(s,\boldsymbol{y}))||_{\mathcal{N}}}{||\eta(\varphi(s,\boldsymbol{x}))||_{\mathcal{N}}}. \tag{A.22}$$

and so it follows that

$$T_{RR}(\boldsymbol{x};\boldsymbol{x}^{\mathrm{ref}}) = T_{RR}(\boldsymbol{y};\boldsymbol{x}^{\mathrm{ref}}) \quad \Leftrightarrow \quad \boldsymbol{x},\, \boldsymbol{y} \in \mathcal{F}_{ss}(\boldsymbol{x}). \tag{A.23}$$

The final statement of Proposition A.7 follows from $\eta(\varphi(s,\boldsymbol{x})) = 0 \Leftrightarrow \varphi(t,\boldsymbol{x}) \in W^{ss}(\boldsymbol{x}^{\mathrm{f}})$ (see Lemma A.9) and the fact that $\eta : B_\kappa(\boldsymbol{x}^{\mathrm{f}}) \to E^s$ is continuous. □

Appendix B

Topology of Sustainable Management

This appendix is taken from [Kit2017e].

B.1 Existence of Eddies

Eddy-like. We call a pair of sets $\mathcal{A}^{+/-} \subseteq \mathfrak{X}$ *eddy-like* iff they fulfills (i) $\mathcal{A}^{+/-} \subseteq \mathfrak{X}^{+/-} - \mathcal{U} - \mathcal{D}$ and (i) $\mathcal{A}^{+/-} \subseteq \text{Capt}\left(\mathcal{A}^{-/+}\right)$. Note the inverted order of the signs in the last term.

Union of two eddy-like pair sets are also eddy-like. We claim that for two eddy-like pairs of sets $\mathfrak{C}_1^{+/-}$ and $\mathfrak{C}_2^{+/-}$ the union pair $\mathfrak{C}_3^{+/-} = \mathfrak{C}_1^{+/-} \cup \mathfrak{C}_2^{+/-}$ is eddie-like, too.

Proof. The first condition is trivially fulfilled and the second one follows straight away from $\text{Capt}\,(\mathcal{A}) \cup \text{Capt}\,(\mathcal{B}) = \text{Capt}\,(\mathcal{A} \cup \mathcal{B})$ for two state sets $\mathcal{A}, \mathcal{B} \subseteq \mathfrak{X}$. □

Existence of eddies. Hence, the union of all eddy-like pairs of sets is maximal and *eddies* exist.

B.2 Parameter Estimation of the AYS-Model

To get a roughly realistic setting, we estimated the parameters of the model using several publicly available data sources.

A_0 was taken from [Cia2013] and slightly rounded. τ^{A} and β were taken from [Kel2011]. ϕ was based on the ton oil equivalent of various fossil fuels and a typical mass share of 90% carbon in fossil fuels, as described in [Nit2016].

Assuming that two degrees warming correspond to a carbon concentration of 450 ppm and thus to a carbon stock of 950 GtC (both being 1.6 times their pre-industrial value), we require that the total growth rate $\beta - \theta A_1$ becomes zero for $A_1 \approx 950\,\text{GtC} - A_0 = 350\,\text{GtC}$, hence θ was taken to be $\beta / A_1 \approx 8.57 \cdot 10^{-5}/(\text{GtC a})$.

ψ was estimated from the World Bank's primary energy intensity data [The2016a].

For τ^{S}, the characteristic depreciation time of renewable energy knowledge, no reliable source was found, so we made a very coarse guess by setting it roughly to the length of an average working life of 50 a.

The break-even knowledge level σ was also estimated very coarsely. According to [Mon2005], past cumulative world consumption of renewable energy is $\approx 2 \cdot 10^{18}$ Btu $\approx 2 \cdot 10^{12}$ GJ or roughly 20 years of world energy consumption. To be on the conservative side and avoid overestimating the potential of renewables, we took σ to be two times that value.

ρ was set as follows. We assume fossil and renewable energy production costs of $C_F \propto F^{1+\gamma}$ and $C_R \propto R^{1+\gamma}/S^{\lambda}$, where $\gamma > 0$ is a convexity parameter and $\lambda > 0$ is a learning exponent. Then energy prices are $\pi_F \propto \partial C_F/\partial F \propto F^{\gamma}$ and $\pi_R \propto \partial C_R/\partial R \propto R^{\gamma}/S^{\lambda}$. In the price equilibrium, $\pi_F = \pi_R$, hence $R/F \propto S^{\lambda/\gamma}$, and thus $\rho = \lambda/\gamma$. According to [Rub2015], the learning rate $LR = 1 - 2^{-\lambda}$ of several renewables is around 1/8, hence $\lambda \approx \log_2(8/7) \approx 0.2$. Assuming a mild convexity of $\gamma \approx 0.1$, we get $\rho \approx 2$.

Appendix C

Comparison of Different Modularity Estimation Algorithms

This appendix is taken from [Kit2017d].

In Figure C.1, we compared five algorithms to estimate the community structure of a network and the resulting modularity: fast greedy [Cla2004], infomap [Ros2008], label propagation [Rag2007], leading eigenvector [New2006b] and WalkTrap [Pon2006]. Details on the choice of WalkTrap are in the caption of Figure C.1.

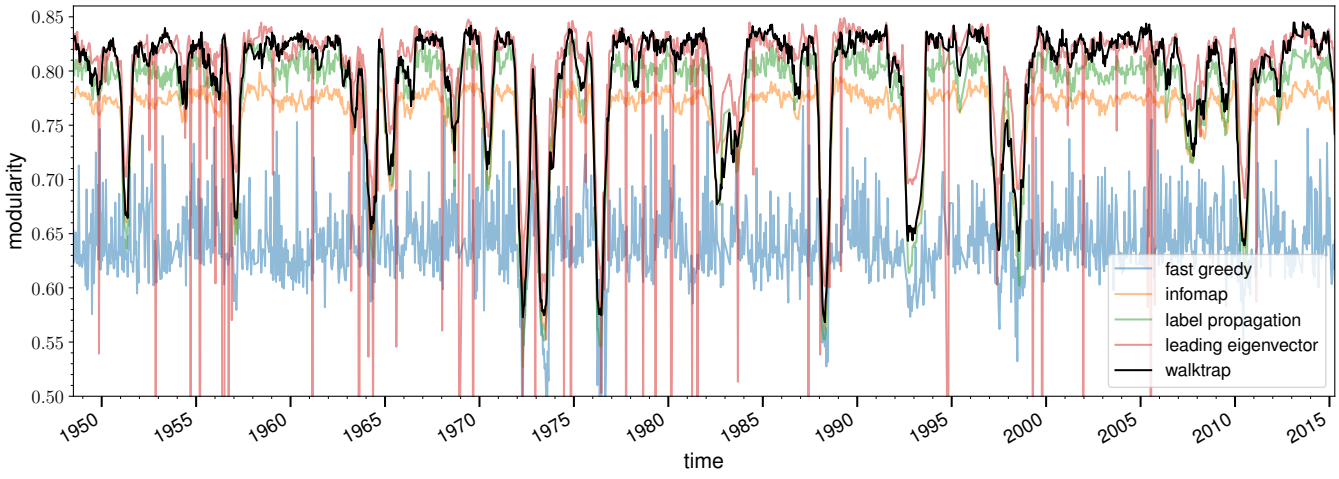


Figure C.1: This figures shows a comparison of five different community detection algorithms for the estimation of modularity: fast greedy, infomap, label propagation, leading eigenvector and WalkTrap. As modularity estimation is a maximization problem, the higher the value the better the result. Visual comparison already shows that leading eigenvector and WalkTrap exceed the others. Furthermore, the leading eigenvector seems to have problem with the specific kind of climate networks we are using here, as there are these strong break-down artifacts. Hence, we decided to use the WalkTrap algorithm. This graphic was published in the appendix of [Kit2017d].

Bibliography

[Alv2016] I. Alvarez, R. Reuillon, and R. D. Aldama. "Viabilititree: a ld-tree Framework for Viability-based Decision". In: *archives-ouvertes.fr* (2016).

[And2013] J. M. Anderies, S. R. Carpenter, W. Steffen, and J. Rockström. "The topology of non-linear global carbon dynamics: From tipping points to planetary boundaries". In: *Environmental Research Letters* 8 (4 2013), p. 044048.

[Ant1971] J. M. Anthonisse. "The rush in a directed graph". In: *Stichting Mathematisch Centrum. Mathematische Besliskunde* BN 9/71 (1971), pp. 1–10.

[Aub1990] J.-P. Aubin and G. Da Prato. "Stochastic viability and invariance". In: *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* 17.4 (1990), pp. 595–613.

[Aub1991] J. Aubin. *Viability theory*. Birkhäuser, 1991.

[Aub1997] J. P. Aubin. *Dynamic economic theory: a viability approach*. Vol. 5. Springer Verlag, 1997.

[Aub2001] J. P. Aubin. "Viability kernels and capture basins of sets under differential inclusions". In: *SIAM Journal on Control and Optimization* 40.3 (2001), pp. 853–881. DOI: 10.1137/S036301290036968X.

[Aub2011] J.-P. Aubin, A. Bayen, and P. Saint-Pierre. *Viability Theory: New Directions*. Springer, 2011.

[Aub2012] J.-P. Aubin and A. Cellina. *Differential inclusions: set-valued maps and viability theory*. Vol. 264. Springer Science & Business Media, 2012.

[Aue2016] S. Auer, K. Kleis, P. Schultz, J. Kurths, and F. Hellmann. "The impact of model detail on power grid resilience measures". In: *The European Physical Journal Special Topics* 225.3 (2016), pp. 609–625.

[Bal2013] G. Balasis, R. V. Donner, S. M. Potirakis, J. Runge, C. Papadimitriou, I. A. Daglis, K. Eftaxias, and J. Kurths. "Statistical Mechanics and Information-Theoretic Perspectives on Complexity in the Earth System". In: *Entropy* 15.11 (2013), pp. 4844–4888. DOI: 10.3390/e15114844.

[Bar1994] G. Barkema, J. Marko, and J. De Boer. "Transient and asymptotic domain growth in the 3D Ising model with conserved spin". In: *Europhysics Letters* 26 (9 1994), p. 653.

[Bar1999] A.-L. Barabási and R. Albert. "Emergence of scaling in random networks". In: *science* 286.5439 (1999), pp. 509–512.

[Bar2001] A.-L. Barabási, E. Ravasz, and T. Vicsek. "Deterministic scale-free networks". In: *Physica A: Statistical Mechanics and its Applications* 299.3 (2001), pp. 559–564.

[Bar2011] R. Barrio, F. Blesa, A. Dena, and S. Serrano. "Qualitative and numerical analysis of the Rössler model: Bifurcations of equilibria". In: *Computers & Mathematics with Applications* 62 (11 2011), pp. 4140–4150.

[Bar2016] W. Barfuss, J. F. Donges, M. Wiedermann, and W. Lucht. "Sustainable use of renewable resources in a stylized social-ecological network model under heterogeneous resource distribution". In: *Earth System Dynamics Discussions* 2016 (2016), pp. 1–13. DOI: 10.5194/esd-2016-15. URL: http://www.earth-syst-dynam-discuss.net/esd-2016-15/.

[Bay2002] Bayen, Crück, and Tomlin. "Guaranteed overapproximations of unsafe sets for continuous and hybrid systems: Solving the hamilton-jacobi equation using viability techniques." In: *Lecture Notes In Computer Science* (2002).

[Bet2016] R. A. Betts, C. D. Jones, J. R. Knight, R. F. Keeling, and J. J. Kennedy. "El Nino and a record CO2 rise". In: *Nature Climate Change* (2016).

[Boc2006] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. "Complex networks: Structure and dynamics". In: *Physics Reports* 424.4 (2006), pp. 175–308. DOI: 10.1016/j.physrep.2005.10.009.

[Bok2006] O. Bokanowski, S. Martin, R. Munos, and H. Zidani. "An anti-diffusive scheme for viability problems". In: *Applied Numerical Mathematics* 56.9 (2006), pp. 1147–1162.

[Bra1988] J. A. Brander and S. M. Taylor. "The simple economics of Easter Island: A Ricardo-Malthus model of renewable resource use". In: *The American Economic Review* 88.1 (1988), pp. 119–138.

[Bri2016] A. Brias, J. Mathias, and G. Deffuant. "Accelerating viability kernel computation with CUDA architecture: application to bycatch fishery management". In: *Computational Management Science* 13.3 (2016), pp. 371–391.

[Car1998] P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre. "Set-Valued Numerical Analysis for Optimal control and Differential Games." In: *Annals of the International Society of Dynamic Games.* (1998).

[Car2011] S. R. Carpenter and E. M. Bennett. "Reconsideration of the planetary boundary for phosphorus". In: *Environmental Research Letters* 6.1 (2011), p. 014009.

[Cas2009] C. Castellano, S. Fortunato, and V. Loreto. "Statistical physics of social dynamics". In: *Reviews of modern physics* 81 (2 2009), p. 591.

[Cha2008] L. Chapel, G. Deffuant, S. Martin, and C. Mullon. "Defining yield policies in a viability approach". In: *Ecological Modelling* 212.1 (2008), pp. 10–15.

[Cha2010] L. Chapel, X. Castelló, C. Bernard, G. Deffuant, V. M. Eguíluz, S. Martin, and M. San Miguel. "Viability and resilience of languages in competition". In: *Plos one* 5.1 (2010), e8681.

[Cha2011] L. Chapel and G. Deffuant. "SVM approximation of value function contours in target hitting problems". In: *Informatics in control, automation and robotics. 8th International Conference, ICINCO 2011*. Vol. 174. Lecture Notes in Electrical Engineering. Springer, 2011, pp. 37–48.

[Cho1979] Y.-C. Chou and W. I. Goldburg. "Phase separation and coalescence in critically quenched isobutyric-acid—water and 2, 6-lutidine—water mixtures". In: *Physical Review A* 20 (5 1979), p. 2105.

[Cho2000] D. Chowdhury, L. Santen, and A. Schadschneider. "Statistical physics of vehicular traffic and some related systems". In: *Physics Reports* 329 (4 2000), pp. 199–329.

[Cho2006] G. H. Choe. *Computational ergodic theory*. Vol. 13. Springer Science & Business Media, 2006.

[Cia2013] P. Ciais, C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, A. Chhabra, R. DeFries, J. Galloway, M. Heimann, C. Jones, C. L. Quéré, R. Myneni, S. Piao, and P. Thornton. "The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change". In: *Change, IPCC Climate* (2013), pp. 465–570. ISSN: 1098-6596. DOI: 10.1017/CBO9781107415324.015. arXiv: arXiv:1011.1669v3.

[Cla2004] A. Clauset, M. E. Newman, and C. Moore. "Finding community structure in very large networks". In: *Physical review E* 70.6 (2004), p. 066111.

[Cla2008] A. J. Clarke. *An introduction to the dynamics of El Niño and the Southern Oscillation*. Academic Press, 2008.

[Cla2011] M. M. Clark. *Transport modeling for environmental engineers and scientists*. John Wiley & Sons, 2011.

[Cru2002] P. J. Crutzen. "Geology of mankind". In: *Nature* 415 (6867 2002), p. 23.

[Cur2005] P. M. Cury, C. Mullon, S. M. Garcia, and L. J. Shannon. "Viability theory for an ecosystem approach to fisheries". In: *ICES Journal of Marine Science: Journal du Conseil* 62.3 (2005), pp. 577–584.

[Cvi2016] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay. *Chaos: Classical and Quantum*. Copenhagen: Niels Bohr Inst., 2016. URL: http://ChaosBook.org/.

[Dai2000] A. Dai and T. M. L. Wigley. "Global patterns of ENSO-induced precipitation". en. In: *Geophysical Research Letters* 27.9 (May 2000), pp. 1283–1286. ISSN: 1944-8007. DOI: 10.1029/1999GL011140. (Visited on 11/25/2016).

[Def2007] G. Deffuant, L. Chapel, and S. Martin. "Approximating viability kernels with support vector machines". In: *IEEE transactions on automatic control* 52.5 (2007), pp. 933–937.

[Def2011] G. Deffuant and N. Gilbert, eds. *Viability and Resilience of Complex Systems: Concepts, Methods and Case Studies from Ecology and Society*. Springer, 2011.

[Del2008] M. Delara and L. Doyen. *Sustainable Management of Natural Resources. Mathematical Models and Methods*. Springer, 2008.

[Dia2001] H. F. Diaz, M. P. Hoerling, and J. K. Eischeid. "ENSO variability, teleconnections and climate change". In: *International Journal of Climatology* 21.15 (2001), pp. 1845–1862.

[Div2017] Division for Sustainable Development Department of Economic and Social Affairs, UNHQ. *Sustainable Development Goals*. Accessed: 2017-11-11. 2017. URL: https://sustainabledevelopment.un.org/sdgs.

[Don2009a] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. "Complex networks in climate dynamics". In: *European Physical Journal Special Topics* 174.1 (2009), pp. 157–179.

[Don2009b] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. "The backbone of the climate network". In: *EPL (Europhysics Letters)* 87.4 (2009), p. 48007.

[Don2011] J. F. Donges, H. C. H. Schultz, N. Marwan, Y. Zou, and J. Kurths. "Investigating the topology of interacting networks–Theory and application to coupled climate subnetworks". In: *European Physical Journal B* 84.4 (2011), pp. 635–651. DOI: 10.1140/epjb/e2011-10795-8. (Visited on 02/18/2015).

[Don2015] J. F. Donges, I. Petrova, A. Loew, N. Marwan, and J. Kurths. "How complex climate networks complement eigen techniques for the statistical analysis of climatological data". In: *Climate Dynamics* 45.9 (2015), pp. 2407–2424. DOI: 10.1007/s00382-015-2479-3.

[Don2017] R. Donner, M. Wiedermann, and J. Donges. "Complex network techniques for climatological data analysis". In: *Nonlinear and Stochastic Climate Dynamics*. Ed. by C. Franzke and T. O'Keane. Cambridge University Press, Cambridge, 2017, pp. 159–183.

[Fan2017] J. Fan, J. Meng, Y. Ashkenazy, S. Havlin, and H. J. Schellnhuber. "Network analysis reveals strongly localized impacts of El Niño". In: *Proceedings of the National Academy of Sciences* 114.29 (2017), pp. 7543–7548. DOI: 10.1073/pnas.1701214114. eprint: http://www.pnas.org/content/114/29/7543.full.pdf. URL: http://www.pnas.org/content/114/29/7543.abstract.

[Fis1989] F. M. Fisher. *Disequilibrium foundations of equilibrium economics*. Cambridge University Press, 1989.

[Fis2005] R. S. Fisher, W. v. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel. "Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)". In: *Epilepsia* 46 (4 2005), pp. 470–472. DOI: 10.1111/j.0013-9580.2005.66104.x.

[Fiu1980] J. Fiutak and J. Mizerski. "Transient behaviour of laser". In: *Zeitschrift für Physik B Condensed Matter* 39 (4 1980), pp. 347–352.

[Fol2010] C. Folke, S. R. Carpenter, B. Walker, M. Scheffer, T. Chapin, and J. Rockström. "Resilience thinking: integrating resilience, adaptability and transformability". In: *Ecology and Society* 15.4 (2010), p. 20.

[Fol2011] C. Folke, Å. Jansson, J. Rockström, P. Olsson, S. R. Carpenter, F. S. Chapin III, A.-S. Crépin, G. Daily, K. Danell, J. Ebbesson, et al. "Reconnecting to the biosphere". In: *Ambio* 40.7 (2011), pp. 719–738. DOI: 10.1007/s13280-011-0184-y.

[For2010] S. Fortunato. "Community detection in graphs". In: *Physics Reports* 486.3 (2010), pp. 75–174.

[Fra1989] H. Frankowska. "Optimal trajectories associated with a solution of the contingent hamilton-jacobi equation." In: *Applied Mathematics and Optimization* (1989), pp. 291–311.

[Fre1977] L. C. Freeman. "A set of measures of centrality based on betweenness". In: *Sociometry* (1977), pp. 35–41.

[Gám2004] A. J. Gámez, C. S. Zhou, A. Timmermann, and J. Kurths. "Nonlinear dimensionality reduction in climate data". In: *Nonlinear Processes in Geophysics* 11.3 (2004), pp. 393–398. DOI: 10.5194/npg-11-393-2004.

[Gau1987] B. D. Gaulin and S. Spooner. "Kinetics of Phase Separation in Mn 0.67 CU 0.33". In: *Physical Review Letters* 58 (6 1987), pp. 668–671.

[Ger2013] D. Gerten, H. Hoff, J. Rockström, J. Jägermeyr, M. Kummu, and A. V. Pastor. "Towards a revised planetary boundary for consumptive freshwater use: role of environmental flow requirements". In: *Current Opinion in Environmental Sustainability* 5.6 (2013), pp. 551–558.

[Gie2007] P. Giesl. *Construction of global Lyapunov functions using radial basis functions*. Springer, 2007.

[Gie2015] P. Giesl and S. Hafstein. "Review on computational methods for Lyapunov functions". In: *Discrete and Continuous Dynamical Systems, Series B* 20 (8 2015), pp. 2291–2331.

[Goz2008] A. Gozolchiani, K. Yamasaki, O. Gazit, and S. Havlin. "Pattern of climate network blinking links follows El Niño events". In: *EPL (Europhysics Letters)* 83.2 (2008), p. 28005. DOI: 10.1209/0295-5075/83/28005. URL: https://doi.org/10.1209/0295-5075/83/28005.

[Har1998] D. Harrison and N. K. Larkin. "El Niño-Southern Oscillation sea surface temperature and wind anomalies, 1946–1993". In: *Reviews of Geophysics* 36.3 (1998), pp. 353–399.

[Has2004] A. Hastings. "Transients: the key to long-term ecological understanding?" In: *Trends in Ecology & Evolution* 19 (1 2004), pp. 39–45.

[Häy2016] T. Häyhä, P. L. Lucas, D. P. van Vuuren, S. E. Cornell, and H. Hoff. "From Planetary Boundaries to national fair shares of the global safe operating space - How can the scales be bridged?" In: *Global Environmental Change* 40 (2016), pp. 60–72. DOI: 10.1016/j.gloenvcha.2016.06.008.

[Hec2016] V. Heck, J. F. Donges, and W. Lucht. "Collateral transgression of planetary boundaries due to climate engineering by terrestrial carbon dioxide removal". In: *Earth System Dynamics* 7 (4 2016), pp. 783–796. DOI: 10.5194/esd-7-783-2016.

[Hei1994] R. Heikes and D. A. Randall. "Numerical Integration of the Shallow-Water Equations on a Twisted Icosahedral Grid. Part I: Basic Design and Results of Tests". In: *Monthly Weather Review* 123.6 (1994), pp. 1862–1880.

[Hei2002] J. Heitzig. "Mappings Between Distance Sets Or Spaces". PhD thesis. Universität Hannover, 2002.

[Hei2012] J. Heitzig, J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. "Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes". In: *The European Physical Journal B-Condensed Matter and Complex Systems* 85.1 (2012), pp. 1–22.

[Hei2016] J. Heitzig, T. Kittel, J. F. Donges, and N. Molkenthin. "Topology of sustainable management of dynamical systems with desirable states: from defining planetary boundaries to safe operating spaces in the Earth system". In: *Earth System Dynamics* 7.1 (2016), pp. 21–50. DOI: 10.5194/esd-7-21-2016. URL: https://www.earth-syst-dynam.net/7/21/2016/.

[Hel2016] F. Hellmann, P. Schultz, C. Grabow, J. Heitzig, and J. Kurths. "Survivability of Deterministic Dynamical Systems". In: *Scientific reports* 6 (29654 2016), pp. 1–12. DOI: 10.1038/srep29654.

[Hir1977] M. W. Hirsch, M. Shub, and C. C. Pugh. *Invariant manifolds*. Vol. 587. Lecture Notes in Mathematics. Springer, 1977.

[Jaf2002] A. B. Jaffe, R. G. Newell, and R. N. Stavins. "Environmental policy and technological change". In: *Environmental and Resource Economics* 22 (2002), pp. 41–69. DOI: 10.1023/A:1015519401088.

[Jol2009] M. Joly and A. Voldoire. "Influence of ENSO on the West African monsoon: temporal aspects and atmospheric processes". In: *Journal of Climate* 22.12 (2009), pp. 3193–3210.

[Kal1996] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph. "The NCEP/NCAR 40-year reanalysis project". In: *Bulletin of the American Meteorological Society* 77.3 (1996), pp. 437–471.

[Kal2012] M. Kalkuhl, O. Edenhofer, and K. Lessmann. "Learning or lock-in: Optimal technology policies to support mitigation". In: *Resource and Energy Economics* 34 (2012), pp. 1–23. DOI: 10.1016/j.reseneeco.2011.08.001.

[Kan2016] A. van Kan, J. Jegminat, J. F. Donges, and J. Kurths. "Constrained basin stability for studying transient phenomena in dynamical systems". In: *Physical Review E* 93 (4 Apr. 2016), pp. 042205-1–7. DOI: 10.1103/PhysRevE.93.042205.

[Kel2011] O. Kellie-Smith and P. M. Cox. "Emergent dynamics of the climate-economy system in the Anthropocene". In: *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* 369.1938 (2011), pp. 868–86. DOI: 10.1098/rsta.2010.0305.

[Kho2017] M. Khodri, T. Izumo, J. Vialard, S. Janicot, C. Cassou, M. Lengaigne, J. Mignot, G. Gastineau, E. Guilyardi, N. Lebas, A. Robock, and M. J. McPhaden. "Tropical explosive volcanic eruptions can trigger El Niño by cooling tropical Africa". In: *Nature Communications* 8 (2017), p. 778. DOI: 10.1038/s41467-017-00755-6.

[Kis2001] R. Kistler, W. Collins, S. Saha, G. White, J. Woollen, E. Kalnay, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino. "The NCEP–NCAR 50–Year Reanalysis: Monthly Means CD–ROM and Documentation". In: *Bulletin of the American Meteorological Society* 82.2 (Feb. 2001), pp. 247–267. DOI: 10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

[Kit2017a] T. Kittel. *Code for reproduction of global teleconnectivity structures of ENSO and large volcanic eruptions*. Accessed 2017-11-07. 2017. URL: https://github.com/timkittel/global-structures-enso-volcanoes/.

[Kit2017b] T. Kittel. *Code for reproduction of the TSM analysis for the AYS model*. Accessed 2017-06-28. 2017. URL: https://github.com/timkittel/ays-model/.

[Kit2017c] T. Kittel. *The PyViability Library*. Accessed 2017-07-18. 2017. URL: https://timkittel.github.io/PyViability/.

[Kit2017d] T. Kittel, C. Ciemer, N. Lotfi, T. Peron, F. Rodrigues, J. Kurths, and R. V. Donner. "Global teleconnectivity structures of the El Niño–Southern Oscillation and large volcanic eruptions: An evolving network perspective". submitted to Nonlinear Processes in Geophysics. 2017.

[Kit2017e] T. Kittel, J. Heitzig, G. Deffuant, J.-D. Mathias, and J. Kurths. "Operationalization of Topology of Sustainable Management to Estimate Qualitatively Different Regions in State Space". submitted to Environmental Modelling & Software. 2017. URL: https://arxiv.org/abs/1706.04542.

[Kit2017f] T. Kittel, J. Heitzig, K. N. Webster, and J. Kurths. "Timing of transients: Quantifying reaching times and transient behavior in complex systems". In: *New Journal of Physics* (2017). URL: http://iopscience.iop.org/10.1088/1367-2630/aa7b61.

[Kli2015] V. V. Klinshov, V. I. Nekorkin, and J. Kurths. "Stability threshold approach for complex dynamical systems". In: *New Journal of Physics* 18 (1 2015), p. 013004.

[Kon2017] C. T. Kone, G. DeSousa, and J. D. Mathias. "Adaptive management of energy consumption, reliability and delay of wireless sensor node: application to IEEE 802.15.4 wireless sensor node". In: *Plos One* (2017), Under revisions.

[Kra2010] P. L. Krapivsky, S. Redner, and E. Ben-Naim. *A kinetic view of statistical physics*. Cambridge University Press, 2010.

[Kuh2010] A. Kuhn and T. Heckelei. "Anthroposphere". In: *Impacts of Global Change on the Hydrological Cycle in West and Northwest Africa*. Springer, 2010, pp. 282–341.

[Kuz1998] Y. A. Kuznetsov. *Elements of applied bifurcation theory, Second Edition*. Vol. 112. 233 Spring Street, New York, NY 10013-1578, USA: Springer Science & Business Media, 1998.

[Lad2015] S. J. Lade, S. Niiranen, J. Hentati-Sundberg, T. Blenckner, W. J. Boonstra, K. Orach, M. F. Quaas, H. Österblom, and M. Schlüter. "An empirical model of the Baltic Sea reveals the importance of social dynamics for ecological regime shifts". In: *Proceedings of the National Academy of Sciences* 112.35 (2015), pp. 11120–11125.

[Len2011] T. M. Lenton. "Early warning of climate tipping points". In: *Nature Climate Change* 1 (4 2011), pp. 201–209.

[Lho2007] Lhommeau, Jaulin, and Hardouin. "Inner and outer approximation of capture basin using interval analysis". In: *4th International Conference on Informatics in Control, Automation and Robotics (ICINCO2007)*. 2007, pp. 1–5.

[Mai2013] J. N. Maidens, S. Kaynama, I. M. Mitchell, M. M. Oishi, and G. A. Dumont. "Lagrangian methods for approximating the viability kernel in high-dimensional systems". In: *Automatica* 49.7 (2013), pp. 2017–2029. DOI: http://dx.doi.org/10.1016/j.automatica.2013.03.020.

[Mar2004] S. Martin. "The cost of restoration as a way of defining resilience: a viability approach applied to a model of lake eutrophication". In: *Ecology and Society* 9.2 (2004).

[Mar2005] D. Maraun and J. Kurths. "Epochs of phase coherence between El Niño/Southern Oscillation and Indian monsoon". In: *Geophysical Research Letters* 32.15 (2005). L15709, n/a–n/a. ISSN: 1944-8007. DOI: 10.1029/2005GL023225. URL: http://dx.doi.org/10.1029/2005GL023225.

[Mat2015] J. Mathias, B. Bonté, T. Cordonnier, and F. DeMorogues. "Using the viability theory for assessing flexibility of forest managers under ecological intensification". In: *Environmental Management* 56 (2015), pp. 1170–1183.

[Mat2017a] J. Mathias, J. Anderies, and M. Janssen. "On our rapidly shrinking capacity to comply with the planetary boundaries on climate change". In: *Scientific Reports - Nature* 7.42061 (2017), pp. 1–7. DOI: 10.1038/srep42061.

[Mat2017b] J. Mathias, S. Lade, and V. Galaz. "Multi-level policies and adaptive social networks - a conceptual modeling study for maintaining a polycentric governance". In: *International Journal of the Commons* 11.1 (2017), pp. 220–247. DOI: 10.18352/ijc.695.

[McC1995] M. P. McCormick, L. W. Thomason, and C. R. Trepte. "Atmospheric effects of the Mt Pinatubo eruption". In: *Nature* 373.6513 (1995), pp. 399–404.

[Men2013] P. J. Menck, J. Heitzig, N. Marwan, and J. Kurths. "How basin stability complements the linear-stability paradigm". In: *Nature Physics* 9 (2 2013), pp. 89–92.

[Men2014] P. J. Menck, J. Heitzig, J. Kurths, and H. Joachim Schellnhuber. "How dead ends undermine power grid stability". In: *Nature Communications* 5 (3969 June 2014), pp. 1–8. DOI: 10.1038/ncomms4969.

[Mit2015] C. Mitra, J. Kurths, and R. V. Donner. "An integrative quantifier of multistability in complex systems based on ecological resilience". In: *Scientific reports* 5 (16196 2015), pp. 1–10. DOI: 10.1038/srep16196.

[Mit2017a] C. Mitra, A. Choudhary, S. Sinha, J. Kurths, and R. V. Donner. "Multiplex-node basin stability in complex dynamical networks". In: *Physical Review E* 95 (3 2017), p. 032317.

[Mit2017b] C. Mitra, T. Kittel, A. Choudhary, J. Kurths, and R. V. Donner. "Recovery time after localized perturbations in complex dynamical networks". In: *New Journal of Physics* 19.10 (2017), p. 103004. URL: http://stacks.iop.org/1367-2630/19/i=10/a=103004.

[Mon2005] Mongabay Organization. *Carbon Dioxide Emissions Charts*. 2005. URL: http://rainforests.mongabay.com/09-carbon_emissions.htm (visited on 12/17/2016).

[Nag2013] B. Nagy, J. D. Farmer, Q. M. Bui, and J. E. Trancik. "Statistical Basis for Predicting Technological Progress". In: *PLoS ONE* 8.2 (2013), pp. 1–7. ISSN: 19326203. DOI: 10.1371/journal.pone.0052669. arXiv: 0508436 [arXiv:submit].

[Nee2003] J. D. Neelin, C. Chou, and H. Su. "Tropical drought regions in global warming and El Niño teleconnections". en. In: *Geophysical Research Letters* 30.24 (Dec. 2003), p. 2275. ISSN: 1944-8007. DOI: 10.1029/2003GL018625. (Visited on 11/25/2016).

[New2004] M. E. J. Newman and M. Girvan. "Finding and evaluating community structure in networks". In: *Physical Review E* 69 (2004), p. 026113. DOI: 10.1103/PhysRevE.69.026113.

[New2006a] M. E. J. Newman. "Modularity and community structure in networks". In: *Proceedings of the National Academy of Sciences* 103.23 (2006), pp. 8577–8582.

[New2006b] M. E. Newman. "Finding community structure in networks using the eigenvectors of matrices". In: *Physical review E* 74.3 (2006), p. 036104.

[New2010] M. E. J. Newman. *Networks, An Introduction*. 1st. Oxford: Oxford University Press, 2010. ISBN: 9780199206650.

[Nit2016] J. Nitzbon. "Bifurcation analysis and parameter estimation in low-dimensional models of global human-nature coevolution". MA thesis. Georg-August-Universität Göttingen, 2016.

[Nit2017] J. Nitzbon, J. Heitzig, and U. Parlitz. "Sustainability, collapse and oscillations of global climate, population and economy in a simple World-Earth model". In: *Environmental Research Letters* 12.7 (2017), pp. 1–15. DOI: 10.1088/1748-9326/aa7581.

[Per2001] L. Perko. *Differential Equations and Dynamical Systems*. 3rd. New York: Springer New York, 2001. ISBN: 0387951164.

[Pon2006] P. Pons and M. Latapy. "Computing communities in large networks using random walks." In: *Journal of Graph Algorithms and Applications* 10.2 (2006), pp. 191–218.

[Rad2013] A. Radebach, R. V. Donner, J. Runge, J. F. Donges, and J. Kurths. "Disentangling different types of El Niño episodes by evolving climate network analysis". In: *Physical Review E* 88.5 (2013), p. 052807.

[Rag2007] U. N. Raghavan, R. Albert, and S. Kumara. "Near linear time algorithm to detect community structures in large-scale networks". In: *Physical review E* 76.3 (2007), p. 036106.

[Ras1982] E. M. Rasmusson and T. H. Carpenter. "Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño". In: *Monthly Weather Review* 110.5 (1982), pp. 354–384.

[Raw2012] K. Raworth. "A Safe and Just Space For Humanity: Can we live within the Doughnut?" In: *Nature* 461 (2012), pp. 1–26. ISSN: 1758-678X. DOI: 10.5822/978-1-61091-458-1.

[Rob2000] A. Robock. "Volcanic eruptions and climate". In: *Reviews of Geophysics* 38.2 (2000), pp. 191–219.

[Roc2009] J. Rockström, W. L. Steffen, K. Noone, Å. Persson, F. S. III Chapin, E. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, et al. "Planetary boundaries: exploring the safe operating space for humanity". In: *Ecology and Society* 14 (2 2009).

[Rop1987] C. F. Ropelewski and M. S. Halpert. "Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation". In: *Monthly Weather Review* 115.8 (Aug. 1987), pp. 1606–1626. ISSN: 0027-0644. DOI: 10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2. (Visited on 11/25/2016).

[Rös1976] O. E. Rössler. "An equation for continuous chaos". In: *Physics Letters A* 57 (5 1976), pp. 397–398.

[Ros2008] M. Rosvall and C. T. Bergstrom. "Maps of random walks on complex networks reveal community structure". In: *Proceedings of the National Academy of Sciences* 105.4 (2008), pp. 1118–1123.

[Rou2013] C. Rougé, J.-D. Mathias, and G. Deffuant. "Extending the viability theory framework of resilience to uncertain dynamics, and application to lake eutrophication". In: *Ecological Indicators* 29 (2013), pp. 420–433. DOI: http://dx.doi.org/10.1016/j.ecolind.2012.12.032.

[Rou2014] C. Rougé, J.-D. Mathias, and G. Deffuant. "Relevance of control theory to design and maintenance problems in time-variant reliability: The case of stochastic viability". In: *Reliability Engineering and System Safety* 132 (2014), pp. 250–260.

[Rub2015] E. S. Rubin, I. M. L. Azevedo, P. Jaramillo, and S. Yeh. "A review of learning rates for electricity supply technologies". In: *Energy Policy* 86 (2015), pp. 198–218. ISSN: 03014215. DOI: 10.1016/j.enpol.2015.06.011.

[Run2012] S. W. Running. "A measurable planetary boundary for the biosphere". In: *science* 337.6101 (2012), pp. 1458–1459.

[Run2014] J. Runge, V. Petoukhov, and J. Kurths. "Quantifying the Strength and Delay of Climatic Interactions: The Ambiguities of Cross Correlation and a Novel Measure Based on Graphical Models". In: *Journal of Climate* 27.2 (2014), pp. 720–739. DOI: 10.1175/JCLI-D-13-00159.1.

[Sai1994] P. Saint-Pierre. "Approximation of the viability kernel". In: *Applied Mathematics and Optimization* 29.2 (1994), pp. 187–209.

[Sai2001] P. Saint-Pierre. "Approximation of viability kernels and capture basins for hybrid systems". In: *Control Conference (ECC), 2001 European*. IEEE. 2001, pp. 2776–2783.

[San1993] B. Sandstede. "Verzweigungtheorie homokliner Verdopplungen". PhD thesis. University of Stuttgart, 1993.

[Sar2007] J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, and J. Kertesz. "Generalizations of the clustering coefficient to weighted complex networks". In: *Physical Review E* 75.2 (2007), p. 027105.

[Sar2010] E. S. Sarachik and M. A. Cane. Cambridge University Press, 2010.

[Sch1993] W. M. Schaffer, B. Kendall, C. W. Tidd, and L. F. Olsen. "Transient periodicity and episodic predictability in biological dynamics". In: *Mathematical Medicine and Biology* 10 (4 1993), pp. 227–247.

[Sch1998] H. J. Schellnhuber. "Discourse: Earth System analysis—The scope of the challenge". In: *Earth System Analysis*. Springer, 1998, pp. 3–195.

[Sch2009] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara. "Early-warning signals for critical transitions". In: *Nature* 461 (7260 2009), pp. 53–59.

[Sch2012] M. Scheffer, S. R. Carpenter, T. M. Lenton, J. Bascompte, W. Brock, V. Dakos, J. Van De Koppel, I. A. Van De Leemput, S. A. Levin, E. H. Van Nes, et al. "Anticipating critical transitions". In: *Science* 338 (6105 2012), pp. 344–348.

[Sch2014] P. Schultz, J. Heitzig, and J. Kurths. "Detours around basin stability in power networks". In: *New Journal of Physics* 16 (2014), p. 125001.

[Ste2010] K. Steinhaeuser, N. V. Chawla, and A. R. Ganguly. "An exploration of climate data using complex networks". In: *ACM SIGKDD Explorations Newsletter* 12.1 (2010), pp. 25–32.

[Ste2011] W. Steffen, Å. Persson, L. Deutsch, J. Zalasiewicz, M. Williams, K. Richardson, C. Crumley, P. Crutzen, C. Folke, L. Gordon, et al. "The anthropocene: From global change to planetary stewardship". In: *Ambio* 40 (7 2011), pp. 739–761. DOI: 10.1007/s13280-011-0185-x.

[Ste2012] K. Steinhaeuser, A. R. Ganguly, and N. V. Chawla. "Multivariate and multiscale dependence in the global climate system revealed through complex networks". In: *Climate Dynamics* 39.3-4 (2012), pp. 889–895.

[Ste2015a] W. Steffen, W. Broadgate, L. Deutsch, O. Gaffney, and C. Ludwig. "The trajectory of the Anthropocene: The Great Acceleration". en. In: *The Anthropocene Review* 2 (1 Jan. 2015), pp. 81–98. DOI: 10.1177/2053019614564785.

[Ste2015b] W. Steffen, K. Richardson, J. Rockström, S. E. Cornell, I. Fetzer, E. M. Bennett, R. Biggs, S. R. Carpenter, W. de Vries, C. A. de Wit, et al. "Planetary boundaries: Guiding human development on a changing planet". In: *Science* 347 (6223 2015), pp. 1259855-1–10.

[Sto2003] H. von Storch and F. W. Zwiers. *Statistical analysis in climate research*. Cambridge University Press, Cambridge, 2003.

[Str1994] S. H. Strogatz. *Nonlinear Dynamics and Chaos*. Perseus Books Publishing, L.L.C., 1994.

[Tan1975] C. Tang, J. Telle, and C. Ghizoni. "Transient effects in wavelength-modulated dye lasers". In: *Applied Physics Letters* 26 (9 1975), pp. 534–537.

[The2016a] The World Bank Group. *Energy intensity level of primary energy*. 2016. URL: http://data.worldbank.org/indicator/EG.EGY.PRIM.PP.KD (visited on 04/11/2016).

[The2016b] The World Bank Group. *GDP*. 2016. URL: http://data.worldbank.org/indicator/NY.GDP.MKTP.CD (visited on 12/17/2016).

[Tre1997] K. E. Trenberth. "The Definition of El Niño". In: *Bulletin of the American Meteorological Society* 78.12 (Dec. 1997), pp. 2771–2777. DOI: 10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2.

[Tre2001] K. E. Trenberth and D. P. Stepaniak. "Indices of el Niño evolution". In: *Journal of Climate* 14.8 (2001), pp. 1697–1701.

[Tso2006] A. A. Tsonis, K. L. Swanson, and P. J. Roebber. "What do networks have to do with climate?" In: *Bulletin of the American Meteorological Society* 87.5 (2006), pp. 585–595.

[Tso2008] A. A. Tsonis and K. L. Swanson. "Topology and Predictability of El Niño and La Niña Networks". In: *Physical Review Letters* 100 (22 June 2008), p. 228502. DOI: 10.1103/PhysRevLett.100.228502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.228502.

[Tso2011] A. A. Tsonis, G. Wang, K. L. Swanson, F. A. Rodrigues, and L. da Fontura Costa. "Community structure and dynamics in climate networks". In: *Climate Dynamics* 37.5-6 (2011), pp. 933–940.

[Tur2004] J. Turner. "The El Niño–Southern Oscillation and Antarctica". In: *International Journal of Climatology* 24.1 (2004), pp. 1–31.

[Van2007] G. Van Geest, H. Coops, M. Scheffer, and E. van Nes. "Long transients near the ghost of a stable state in eutrophic shallow lakes with fluctuating water levels". In: *Ecosystems* 10 (1 2007), pp. 37–47.

[Wat2016] C. N. Waters, J. Zalasiewicz, C. Summerhayes, A. D. Barnosky, C. Poirier, A. Gałuszka, A. Cearreta, M. Edgeworth, E. C. Ellis, M. Ellis, et al. "The Anthropocene is functionally and stratigraphically distinct from the Holocene". In: *Science* 351 (6269 2016), aad2622-1–10. DOI: 10.1126/science.aad2622.

[Wec2013] T. Weckesser, H. Jóhannsson, and J. Østergaard. "Impact of model detail of synchronous machines on real-time transient stability assessment". In: *Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid, IREP Symposium*. IEEE. 2013, pp. 1–9.

[Wie2013] M. Wiedermann, J. F. Donges, J. Heitzig, and J. Kurths. "Node-weighted interacting network measures improve the representation of real-world complex systems". In: *EPL (Europhysics Letters)* 102.2 (2013), p. 28007.

[Wie2015] M. Wiedermann, J. F. Donges, J. Heitzig, W. Lucht, and J. Kurths. "Macroscopic description of complex adaptive networks coevolving with dynamic node states". In: *Physical Review E* 91.5 (2015), p. 052801.

[Wie2016] M. Wiedermann, A. Radebach, J. F. Donges, J. Kurths, and R. V. Donner. "A climate network-based index to discriminate different types of El Niño and La Niña". In: *Geophysical Research Letters* 43.13 (2016), pp. 7176–7185. ISSN: 19448007. DOI: 10.1002/2016GL069119. arXiv: 1604.04432.

[Wik2017] Wikimedia Commons. *Attractor*. 2017. URL: https://en.wikipedia.org/wiki/Attractor.

[Yam2008] K. Yamasaki, A. Gozolchiani, and S. Havlin. "Climate Networks around the Globe are Significantly Affected by El Niño". In: *Phys. Rev. Lett.* 100 (22 June 2008), p. 228501. DOI: 10.1103/PhysRevLett.100.228501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.228501.

[Yua2003] Y. Yuan, J. Kubokawa, and H. Sasaki. "A solution of optimal power flow with multicontingency transient stability constraints". In: *IEEE Transactions on Power Systems* 18 (3 2003), pp. 1094–1102.

[Zgl1997] P. Zgliczynski. "Computer assisted proof of chaos in the Rössler equations and in the Hénon map". In: *Nonlinearity* 10 (1 1997), p. 243.

Selbstständigkeitserklärung

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 126/2014 am 18.11.2014 angegebenen Hilfsmittel angefertigt habe und alle Quellen angegeben habe.

_____ _____

Place, Date Signature